

nature

5 October 1978

Europe should invest more in space science

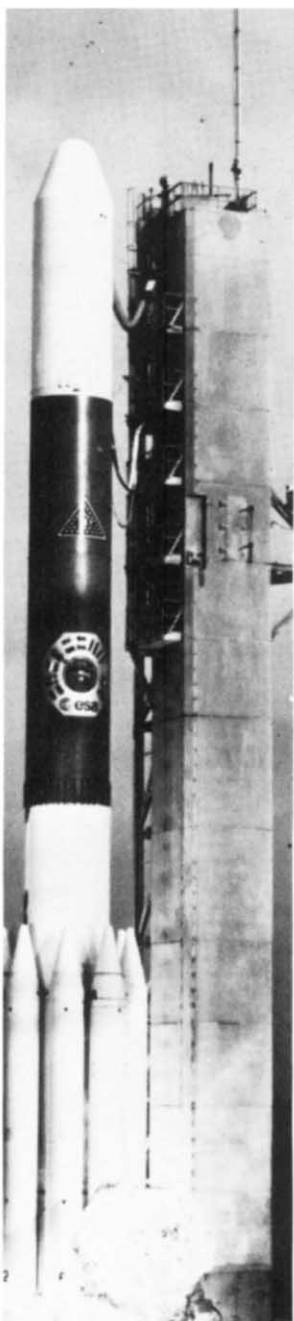
SPACE science is expensive. NASA's science budget amounts to half-a-billion dollars a year. No individual European country could afford a national programme to satisfy the needs of its astrophysicists and geophysicists to do experiments in space, let alone provide opportunities for scientists working in other fields to realise the potential space has to offer. For Europeans, therefore, international collaboration is essential. It pools funds, expertise and ideas.

The European Space Agency (ESA) undertakes this task. If ESA did not exist, space scientists in individual European countries would have to rely on *ad hoc* international collaborations. More likely than not, most of these would involve NASA, which has a world lead in space science and can build spacecraft more cheaply and efficiently than any European country. Such collaborations could undoubtedly benefit all taking part: previous ones have proved this.

But total reliance on NASA could be dangerous. For example, some European scientists might want to build a mission which is of no interest to NASA at the time it is proposed. This might not be because it is unworthy of development but simply because it is not NASA's policy to build that type of mission at that time. European scientists need some independence and ESA should be able to provide it.

This is not to say that ESA itself should shun collaboration with NASA. On the contrary, wherever such collaboration is the best way of achieving the scientific aims of a particular project, it should be done. And this applies just as much to collaboration between individual countries and NASA.

If ESA is to provide European scientists with a certain amount of independence, it must provide them with enough launches. Many European scientists feel that, because of its appallingly low budget for science—about a tenth that of NASA—and poor value for money, ESA is failing them. And matters are unlikely to improve. The few funds which are available might well be spread even more thinly in future.



ESA is now under pressure to support missions to the planets. And it will soon have to decide how to pay for materials and sources, such as climatology.

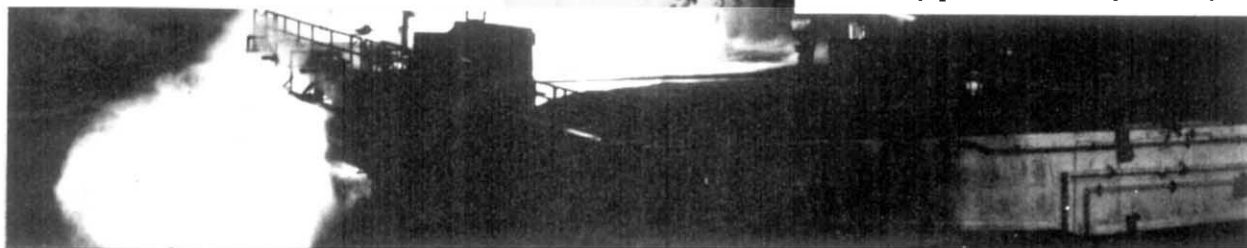
Although it cannot support some of these scientific disciplines at the expense of the others, ESA must make sure that it does not lose the support of its best scientists. There is a danger that when a group of space scientists in Europe develops a world lead in its discipline, it will want to break away from ESA because ESA cannot give it the opportunities it needs to maintain that lead.

The key to ESA's future success is more space opportunities more often. This could be achieved if member states agreed to increase the mandatory science budget. Matters could be improved even further if ESA could cut costs.

ESA's Science Advisory Committee is arguing that some of the surplus money which will be available in the early 1980s—when the applications programmes start to cost less (see page 356)—should be diverted to science. But space science is not high on the list of priorities in most member states, who want an immediate financial return for their investments.

Good science, as always, has its intrinsic merit, but there are extrinsic reasons for a substantial investment in space science. In many industries, a strong programme of scientific research stimulates the development of new technologies. This applies *a fortiori* to the space industry. Building spacecraft for scientific research often calls for greater technological advances than building applications satellites. For example, ESA's faint object camera for the Space Telescope has put Europe's industry in the forefront of advanced photon detector technology.

Further, if the development and running of space applications is taken over from ESA by industry, as seems possible, further investment in space science will be the one way in which national governments can keep their space industries on their toes. Space science should not be seen as a luxury, but as a necessary part of a healthy industry. □



INSIDE: eight-page special on the European space science programme



Scientists press ESA to restore budget to 1971 levels

ESA's future is in the balance. The end of the decade set aside for developing space applications is in sight. At the end of this year, the ESA Council will have to make some fundamental decision on the future.

Exactly how scientists should influence that future was discussed last week at a meeting of ESA's Science Advisory Committee (SAC) in Nice. "Now is the time to say that science needs more support from the member states", is how Professor Klaus Pinkau, a member of the SAC summed up the conclusion of the meeting.

The SAC would like to see some of the funds now spent on applications diverted to science in the future. In 1971, ESA's science budget was cut to make room for the new space applications. Now that many of those applications are nearing the end of their development there is a strong case for

some of the money which will be released to be spent on science.

The other factor which distinguishes ESA's present predicament from the past is how to incorporate the 'new' fields of materials and life sciences and earth resources into the science programme. It is essential that the new fields should at least be screened for their scientific worthiness by ESA's scientific advisory bodies. If they are funded outside the mandatory science budget the situation could arise where a project in one of the 'new' fields would be funded, even though scientifically it was not as worthy as a geophysics or astrophysics experiment, simply because a separate budget had been set aside for it.

The West German government is likely to support the scientists' view. It has already said that it would like to divert funds from applications to



science. The French government, however, might need some hard persuasion from its scientists. And in the UK, diverting funds from industry to science is likely to produce many administrative difficulties because of the way it pays its ESA subscriptions.

The SAC will present its ideas to the Science Programme Committee (SPC) in November. At the end of the year, it will report to the ESA Council outlining its ideas for the future. □

What the critics say

EUROPEAN space scientists are far from content with the service they receive through the European Space Agency (ESA). According to Professor Martin Rees, who until last year was chairman of ESA's Science Advisory Committee, ESA's science budget is only about a tenth of NASA's. And many claim that ESA spends it inefficiently.

ESA was created in 1975 out of the European Space Research Organisation (ESRO) and the European Launcher Development Organisation (ELDO). The former had been solely responsible for space science: the latter for developing an independent launcher capacity for Europe. Since these two objectives have been under the umbrella of one agency, there have been fears that too much attention is being paid to the industrial programme (it takes 85% of ESA's budget) at the expense of science. According to one West German official "science developed a good reputation in ESRO, but ESA is not taking enough care of it". In 1971 ESRO defined a 'minimum viable budget' which amounted to only two-thirds of its expenditure, and ESA has kept to it since.

The scientists' dissatisfaction is matched by the frustration of some of ESA's staff who feel that ESA is blamed for many ills for which it cannot be held entirely responsible. They have the exceedingly difficult task of

trying to implement a European science programme within the constraints of a very limited budget (about £50m per year, compared to CERN's £190m for high energy physics). And in recent years the reduction in many national programmes has put an even greater burden on that budget.

So far all ESA's scientific missions have been paid for out of the mandatory science budget. Each of the 11 member states (Belgium, Denmark, France, Germany, Ireland, Italy, the Netherlands, Spain, Sweden, Switzerland, and the UK) contributes to the mandatory budget in proportion to its gross national product up to a maximum of 25%.

But as time goes on, European space scientists need larger space missions more often because they need to build on the previous research of European and national missions. On top of that, scientists in fields which have never been catered for by a European space mission, in particular planetary science and climatology, are putting pressure on ESA to provide them with opportunities. And looking further to the future materials and life scientists might well be applying similar pressure if they gain their entrée into space through the launch of a manned laboratory—Spacelab in 1983. According to one West German scientist "a multiplicity of disciplines are now knocking on the mandatory science budget's door".

The difficult position ESA is in is illustrated in the long term planning

reports produced by its astronomy and solar systems working groups in 1976. They both concluded that a basic programme in each of their respective fields over the next ten years would take up the entire mandatory budget available over that time.

It is also claimed that ESA is inefficient. In recent evidence to the Committee on Science and Technology, US House of Representatives, Dr Wolfgang Finke, speaking in his capacity as Head of Space and Transportation of the Bundesministerium für Forschung und Technologie (BMFT), said that a conservative estimate of the cost of an international project "is more or less equal to the product of the 'true' costs of the project done nationally and the square root of the number of participating countries". For ESA with eleven member states this rule implies very expensive projects indeed. However, a direct comparison of the costs of ESA's satellites and NASA's or those built in an individual country, is very difficult because of differences in accounting and charging overheads. Nevertheless ESA is in the process of making such a comparison.

A major factor contributing to high costs is ESA's industrial policy. The policy was created to help develop the European aerospace industry and states that industrial contracts have to be distributed throughout member states in proportion to the size of their contributions to ESA's subscription. Thus an ESA satellite may be built in

*This report was compiled
by Judy Redfearn*

several firms in several member states. There is little or no competition for contracts and therefore little natural price control. Also because space contracts are few and far between, and are only small fry to the aerospace industry, they suffer from lack of continuity and standardisation.

Standardisation could only be introduced if ESA built a series of similar satellites. However, it is unlikely that enough satellites could be built with the available funds to make standardisation a reality. In the early days, ESRO attempted to build a standard bus, but after the first, TDI, a uv satellite launched in 1972, it had to scrap the idea. "European missions have never been suitable for a standard bus" says Mr Roy Gibson, ESA's Director-General. Nevertheless, Dr Edgar Page, Head of the Space Science Department at ESA's technology centre, (ESTEC), in Noordwijk the Netherlands, believes that ESA's future success will depend to a large extent on whether or not it can achieve some sort of standardisation. At the moment, ESTEC engineers are looking at standard subsystems rather than standard buses, but these are more likely to find uses in applications satellites rather than scientific ones.

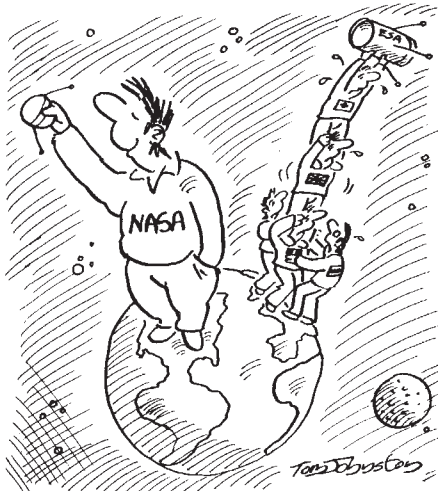
Reusing spacecraft might be another way of cutting costs. At the moment NASA's Space Shuttle is the only vehicle capable of retrieving or refurbishing spacecraft in orbit, so the cost-effectiveness of doing this will be very dependent on its launch costs and precisely how it is used. So far ESA is participating in two future reusable missions: Spacelab and NASA's Space Telescope. Spacelab, however, will probably be very expensive to operate, one of the reasons being the high cost of using the Shuttle. NASA is quoting \$31m for a Space Shuttle launch by non-US government users.

Some of the missions ESA has already flown and many of the ones it is considering for the future are too large and expensive for any individual country to build. Mr Ernst Trendelenburg, ESA's Director of Scientific Programmes, believes that these are precisely the kind of missions ESA should be adopting. Mr Gibson, however, believes that ESA should also provide smaller facilities. These differing views are reflected in the scientific community itself. Largely, big launches mean high energy astronomy, and a rather small community of users, and small launches mean geophysics and more users, except for certain very sophisticated experiments.

Until 1972, scientists needing small payloads carrying experiments of short duration were catered for by ESA's sounding rocket programme. This was stopped, however, on the grounds that

sounding rockets are within the scope of national programmes. Now ESA only helps out in the running of ESRANGE, a rocket range in Sweden. The main hope for the future in catering for small payloads is Spacelab. But there are fears that it might be too expensive to be used to best effect. "We had hoped Spacelab would be a seven day balloon or rocket facility", commented one German scientist "but it looks like becoming a seven day satellite facility". And for atmospheric scientists, at least, it will not be as useful as sounding rockets because it pollutes its environment and its orbit is uninteresting to them scientifically.

ESA's scientific satellites have become more complex as time has passed. In astrophysics, for example, early satellites such as TDI were designed to survey the sky at various wavelengths and identify interesting objects. Follow-up satellites need to observe those objects with greater pointing accuracy and higher resolution. Large



observatories in space are needed to do this. The instruments for them often have to be built under the direction of ESTEC because they are beyond the technological capabilities of universities and research institutes. Many scientists feel that this trend is taking away their opportunities to design and build instruments which are just as important for maintaining a strong scientific community as being able to do theoretical work from satellite data. ESA is developing a will of its own.

One possibility for overcoming this problem is to build large observatories within consortia of institutes. Professor R. Lust of the Max Planck Society in Munich and the first scientific director of ESRO, believes that this is possible. "EXOSAT (ESA's X-ray satellite to be launched in 1981 on Ariane) is a large observatory satellite and it is being built in industry under the direction of ESTEC. Cos B (ESA's cosmic ray satellite launched in 1975) was just as complicated and it was

built successfully in five different institutes".

Not all scientists are so critical of ESA. In spite of its shortcomings some think it provides an essential service. Professor van de Hulst of Leiden University, the Netherlands, for example believes that "ESA suffers from the problems of a big international organisation but it must be judged on the projects it has done. These have been good." And Professor Roger Bonnet, of the Laboratoire de Physique Stellaire et Planetaire, near Paris and chairman of the SAC thinks that it offers a certain independence for Europe. "Cooperation with NASA can be difficult", he says "because NASA won't always be willing to adapt its programmes."

Inevitably, the least discontented scientists are those in countries with an active national programme. Holland, for example, has a small but world-renowned astrophysics community which has been able to take advantage of opportunities on some ESA satellites, such as Cos B, and on other satellites in collaboration with other countries such as the NASA-Dutch X-ray satellite ANS-1 and the NASA-Dutch-UK infrared satellite, IRAS due for launch in 1981. It is just beginning to consider what kind of satellite to build after IRAS. Two possibilities are an X-ray satellite and another IR satellite.

Perhaps the country where ESA's problems are posed most acutely is the UK. Unlike most other European countries, its contributions are paid by two agencies out of two separate budgets. The Science Research Council (SRC), which pays about £8m, is responsible for science and the Department of Industry, which pays about £30m, for applications. The SRC therefore is a focus for scientific opinion within the UK government. In 1975, during a period of financial stringency, the SRC adopted the policy of squeezing funds for big science. Before then the UK's national space programme exceeded its contribution to ESA. Now it is slightly less and the current national programme will come to an end next year after the launch of UK-6, an X-ray satellite costing some £9m. Apart from the IRAS satellite, to which the UK is making a small contribution, the UK has no other approved international collaboration outside ESA. Even the sounding rocket programme has been cut to the bare bones. With so little money, the SRC is concerned about the value for money it is getting with ESA.

It is the British astrophysics community which is feeling the most left out. In particular, UK X-ray astronomers—who have not been well served by ESRO and ESA—are looking for

further opportunities to help them maintain the world lead they gained after UK-5, an X-ray satellite launched in 1975. They feel that the data from EXOSAT, ESA's large X-ray observatory, will be spread on the ground too thinly to be of much use to them. "The future of British astrophysics does not lie with ESA" commented one British high energy astronomer.

Recently the SRC has been devoting some thought to ways of providing UK scientists with more space opportunities and of cutting the cost of space science. It is considering collaborating with NASA on a multi-disciplinary refurbishable satellite (MRS) to be launched and refurbished by NASA's Space Shuttle. The idea is to increase the lifetime and usefulness of a single satellite by repairing it and altering its payload either in orbit or on the ground. The SRC hopes that an MRS could be a very cheap way of doing both astrophysics and geophysics experiments. There are some fears, however, especially among geophysicists that by trying to please everyone it will please no one.

The SRC would like to involve ESA in studies on the MRS, but enthusiasm for the idea within ESA is low. There are doubts about its feasibility and worries that it would mean yet another collaboration with NASA. Nevertheless, ESA will at least be obliged to look into the idea although it could not be studied in time for ESA's next project decision at the end of 1979.

West Germany's policy is to conduct as much of its space science through ESA as possible and it would actually like to increase its contribution to ESA's mandatory science budget. However, it could only do this if all member states agreed to increase their contributions accordingly. In the long term, Germany would like to see funds diverted to science from applications. "Applications should only go on in ESA for a short time," says Dr H. Strub of the BMFT, "then they must be marketed and taken out of ESA's control. In science ESA runs and develops satellites but in applications, the users must take this over once the technology has got off the ground". In particular, Germany would like to see more spent on science when ESA's own launcher, Ariane, and Spacelab near the end of the most expensive phases of their development. By the mid-1980s, it would like ESA's entire budget to level off at 350 mau (about £230m) at today's prices (it is now about 530 mau, £345m) and most of it to be spent on science and new fields to ESA bordering on applications and science such as remote sensing and climatology.

The greatest pressure on ESA to do planetary science comes from

Germany. German scientists would like to consolidate the experience they gained from the Helios probes, two solar physics missions launched in 1974 and 1976 in collaboration with NASA. The French would also like to build on their collaborations with the Soviet Union on a number of planetary missions. Germany is also considering collaborating with NASA on Robi, a cheap X-ray satellite.

Scientists within the SAC feel many of ESA's problems could be solved if member states agreed to increase the mandatory science budget and ESA could increase its cost-effectiveness. Professor Klaus Pinkau of the Max Planck Institute für Extraterrestrische Physik recently sent out a questionnaire to European space scientists canvassing opinion on ESA. "The science budget doesn't need a six-fold increase" Professor Pinkau told *Nature*. "This would be out of proportion to the funding of the rest of science". He would like to see the mandatory science budget restored to its pre-1971 level, about 100 mau (£65m) at today's prices, and would like ESA to make efforts to increase its cost-effectiveness by at least 20%. Over a ten year period, he says, ESA

should be able to fly one mission a year if it built five missions costing 50 mau each, four costing 120 mau and one very good and expensive one costing 250 mau.

To achieve this, however, ESA would have to make special efforts to keep the cost of small satellites within the 50 mau limit. Professor Pinkau suggests that facilities on the ground could be used more efficiently so that the science budget does not have to pay for facilities which are not being used for some part of the year. "Good management should show where you lose money", he says. The industrial policy might also have to be revised so that the principle of fair return would apply over say ten years.

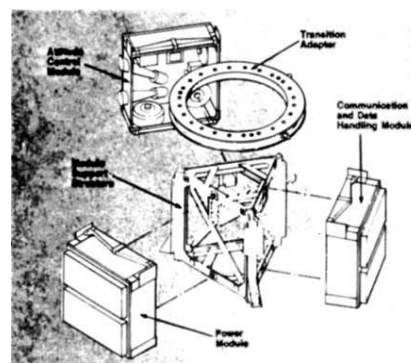
This scheme, he feels could incorporate materials and life sciences into the mandatory budget. Earth resources would have to remain outside the budget but could still be refereed within the procedure of the SAC. Such a scheme however is dependent on the member states agreeing to increase the science budget and on ESA achieving a substantial reduction in costs. The former lies within the power of national governments. The latter is up to ESA. □

The multi-discipline refurbishable satellite

THE MRS is one country's (the UK's) attempt to find more flying time outside the auspices of ESA. Under consideration by the Astronomy, Space and Radio Board of the Science Research Council, it would use The Godard Space Flight Center's kit-form Multi-mission Modular Spacecraft (MMS) as a base (see figure) and erect on it instruments to serve both astronomers and geophysicists. A study is underway (it will report in January) at British Aerospace, Stevenage, to assess the difficulties of combining the instruments on the same vehicle. "And it doesn't make life easy" said one scientist.

The MRS would be a collaboration with NASA, which is interested in flying high energy astronomy experiments. But NASA is not concerned with flying geophysics on this mission; so geophysicists in the UK are suspicious that the exercise is one designed by high energy astronomers to 'sell' the project to the Science Research Council, which is conscious of the need to serve as many scientists as possible.

Funds would be needed by about the end of 1980, for a launch from the Space Shuttle in 1985. The UK would assemble the satellite, at a cost of £15-£20m—twice the cost of UK6, for example. But the MRS could take a payload of up to 1,000 kg, and could



Exploded view of MMS basic modules and elements

have a continuous data rate an order of magnitude greater than UK6.

Atmospheric physics may be the casualty of the multidiscipline confrontation, as its requirements are least compatible with astronomy. But the MRS is retrievable and refurbishable and is designed to make use of the supposed reduced costs of reuse of the MMS base and some instrumentation. Geophysics could get a better look in on a subsequent launch.

But everything depends on whether an agreement can be reached with NASA on the first and subsequent flights—as always with such collaborations—and, ultimately, on the uncertain cost of operations with the shuttle. □

THE ORGANISATION



Roy
Gibson
ESA's
Director-
General

Who makes the decisions?

FINAL approval for ESA's scientific projects is granted by the Science Programme Committee (SPC), the body responsible to the ESA Council for running the science programme within the financial limits set by the Council. Each member state is represented on the SPC by its own delegates and is allowed one vote in any decision.

Projects for final selection are put to the SPC by the Director-General, Mr Roy Gibson. He is advised by the Science Advisory Committee (SAC) and four working groups: the astronomy working group (AWG), the solar system working group (SSWG), the materials sciences working group (MSWG) and the life sciences working group (LSWG). Experts appointed by the Director-General and acting in their own personal capacity, sit on each advisory body. The MSWG and LSWG were recently created. The two areas of science they serve are new to ESA. One life scientist, one meteorologist and four representatives of the astrophysical and geophysical sciences sit on

the SAC. Its meetings are attended by the chairmen of the working groups.

The executive recommends that new project proposals are screened by the working groups. The money spent on early studies is authorised by the Director-General.

If preliminary studies are successful, the Director-General may recommend that a project be considered by the SPC. The SPC will then decide whether or not to go ahead. When a project has been approved, ESA puts out a call for proposals for experiments.

Although there is no formal link between the working groups and the SAC, in practice they work together quite closely. Typically, out of 10-20 project proposals put before the working groups, 5-6 will be passed on to the SAC and three on to the SPC. On average the SPC chooses one major project for development every one and a half to two years.

So far, the SPC has only approved one experiment in the materials and life sciences, the Sled for Spacelab, to be funded out of the mandatory science budget. The rest of the experiments approved in these disciplines are for the first Spacelab payload and are being paid for in individual countries and out of the Spacelab budget.

As well as vetting individual project proposals, the working groups and the SAC discuss overall policy for the science programme. In 1976, for example, the AWG and SSWG both produced reports on the science and projects needed in their respective fields up to 1990. About once every four years, the SAC also discusses plans for the future, such as it did last week.

Responsibility for seeing that the science programme is put into practice

rests with Mr Ernst Trendelenburg, Director of Scientific Programmes. His responsibility now includes future science programmes. To speed up decision-making, he would like to see the number of early studies cut, especially those for projects which are obviously unsuitable for further development from the outset. □

Four steps to get a project off the ground

A SCIENTIFIC mission proposal which has been studied by the relevant working group and found scientifically worthy will go through four 'stages' before it is finally approved and the spacecraft is built. Each stage is independent of the others so that the mission can be scrapped at the end of any one without breaking contracts or jeopardising the money which would be spent in the subsequent stages. Most projects which go beyond phase A, however, are completed. The steps in a mission's development are as follows:

Mission definition or assessment. This is a study stage to define the mission and its payload. It is done under the direction of ESA by seven to ten scientific experts in the mission's field of study. They measure the interest shown in the mission by the scientific community in Europe, assess its scientific soundness and the ability of European universities and institutes to build instruments capable of achieving its scientific objectives. They also predict what sort of system will evolve, how

● turn to next page

The cost: where the money comes from, where it goes

ESA's TOTAL budget for 1978 is about 530 mau (one mau is roughly US \$1.3 million). Of this, 75.5 mau is to be spent on the science programme, about 80 mau on the general budget and the rest on Spacelab, Earthnet and applications programmes such as telecommunications satellites, Meteosat and the development of ESA's own launcher, Ariane.

The science budget and part of the general budget are mandatory. Each member state contributes to them in proportion to its gross national product. Thus, in 1976, Belgium contributed 3.48%, Denmark 2.26%, France 21.20%, Germany 25.25%, Italy 13.23%, the Netherlands 5.34%, Spain 4.44%, Sweden 4.46%, Switzerland 3.56% and the United Kingdom 16.47%. So far all ESA's scientific missions have been conducted through the mandatory programme, even

though, in principle, nothing excludes a group of interested member states from funding an optional one.

The mandatory part of the general budget funds 'basic activities' such as the technological research programme at ESTEC, the salaries of some headquarter's staff, the documentation centre at Frascati and the infrastructure. In 1978 it amounted to about 54 mau. The rest of the general budget is for facilities which cannot be charged to any particular programme such as the cost of buying the agency's headquarters in Paris and the cost of running the Ariane's launch pad at Kouru in French Guiana.

All of the science budget is spent on building spacecraft and very large instruments such as those for Spacelab or large observatory-type free-flyers, and on maintaining facilities for tracking them and receiving data. Most of

the experiments on board satellites, however, are built in universities and institutes and are funded out of national funds for scientific research.

In 1971 the mandatory science budget was fixed at 28 mau. In 1978 that level is 75.5 mau and in 1979 it is expected to be about 79 mau. As scientific satellites cost between 50-100 mau, and some of them even more, this level of funding means that ESA can build one every one and a half to two years.

The capital cost of launches is not the only expense, and already the science budget is almost fully accounted for until the end of 1980. It is proving difficult to find the 2 mau needed to keep Cos B, ESA's cosmic ray satellite launched in 1975, operating for another year. This sum has not already been included in next year's budget because Cos B has exceeded its expected lifetime. □

much it will cost and what sort of effort in terms of manpower and technological research will be needed. The study normally takes three months.

Phase A or study stage. Those proposals not thrown out after the assessment stage are subjected to a more detailed definition. This is usually done under the direction of a technical study team which may contract out to industry some parts of the study. It involves a detailed estimate of the cost and time to completion and a thorough description of the payload, spacecraft, orbit, ground support and data handling and analysis. Typically the study stage lasts nine months and costs 1% of the final project cost.

Phase B or design stage. A scientific project which enters this stage has usually been approved for development by the SPC. It is put under the management of a project team which will be responsible for it until it has been launched and is operating. The project team places contracts in industry for the design of the spacecraft. Typically, phase B takes one year and costs 10% of the final cost of the project.

Phase C/D or development and production. Typically, for a scientific satellite, this will take three years and cost 40–70 mau.

The project teams include engineers from ESTEC's department of development and technology (DDT). When they are not part of a project team they do research in one of the DDT's divisions on various aspects of advanced space technology. This research is important for keeping Europe in the forefront of space technology.

Before launch all ESA satellites are tested at ESTEC. It has the only facility in Europe—a high vacuum chamber 14 m high and 10 m in diameter—large enough to test a satellite with its booms deployed or the nose cone of Ariane. Once the satellite is in orbit, however, operational responsibility for it shifts to the European Space Operations Centre (ESOC) at Darmstadt, West Germany. As well as control of the orbit, ESOC is responsible for setting up ground stations to receive data.

One of ESA's objectives is to encourage development of European industry. This is reflected in the way it puts out tenders to industry for its scientific satellites. Its industrial policy states that each member state should receive the same percentage of the total contract put out to tender as it contributes to the mandatory science budget. This is irrespective of whether it is participating in the scientific payload. This policy is controlled by the Industrial Policy Committee (IPC) which can suggest how contracts should be placed to ensure that each country gets a fair return. □

ESA's SCIENTIFIC RECORD

The past . . .

ESA and its predecessor ESRO have launched twelve scientific satellites since the first launch in May 1968. Six of them are still operating: Cos B, a highly successful gamma ray observatory launched in April 1975; Geos 1, which failed to reach geostationary orbit in April 1977 but is producing some useful magnetospheric data; ISEE-2, which was launched in October 1977 and is observing the interaction of the solar wind and the Earth's magnetosphere in conjunction with NASA's ISEE-1; Meteosat, launched in November 1977, contributing to the Global Atmospheric Research Programme and the World Weather Watch; International Ultraviolet Explorer, the ultraviolet astronomical observatory, launched as a NASA-UK-ESA collaboration in January 1978; and Geos 2, a 'copy' of Geos 1, successfully launched into geostationary orbit in June 1978 to observe the magnetosphere.

Of the previous seven satellites, five were designed for magnetospheric and atmospheric experiments, and two for high energy astronomy: ESRO-II for cosmic rays and solar X-rays, launched May 1968) and TD-1 (a UV satellite launched March 1972 and operational for two years but troubled by malfunctioning tape recorders). Of the complete list of thirteen satellites, nine were for Earth-related observations (mostly on Sun-Earth interaction) and four for astronomy. Astronomical satellites were launched in 1968, 1972, 1975 and 1978.

The present . . .

ESA's report to the 21st COSPAR meeting in June 1978 says that while Sun-Earth relationships "continue to be an important feature of the agency's scientific programme", there is a change afoot: "an increased emphasis is being placed on astronomical studies". This is reflected to some extent in the current list of approvals (four main projects, two of which are astronomical), and to a greater extent among the competing proposals for the next project (three out of five are for astronomy). Approved are:

The first Spacelab payload. Spacelab is scheduled to fly with the tenth flight of the Space Shuttle on 3 December 1983. It will be manned, but will stay aloft for only 7–30 days. According to ESA's 1978 report to COSPAR "the composition of a typical Spacelab payload will be distinctly different from

that of a conventional research satellite". Life sciences and materials sciences will benefit from 'look-see' experiments in a new environment. But the observing time is too short for much astronomy.

EXOSAT. This 0.1 keV to 50 keV X-ray astronomy satellite is scheduled for launch on Ariane in February or March 1981 (or by Thor-Delta if Ariane is not ready). It has two operating modes: occultation, using the Earth or Moon to determine the position and structure of X-ray sources, and arbitrary pointing, which will be used to determine this spectral and temporal variation.

Out of Ecliptic/Solar Polar Mission.

This is a joint mission with NASA to launch two spacecraft in February 1983, to fly out of the plane of planetary orbits and around the North and South poles of the Sun. The gravitational field of Jupiter will be used to swing the spacecraft out of the ecliptic. The objective is to extend into three dimensions our knowledge of the solar corona, the solar wind, the solar magnetic field, the cosmic ray flux, and interplanetary/interstellar neutral gas and dust.

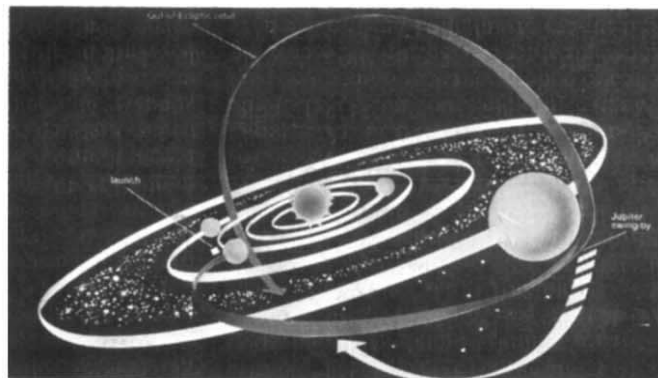
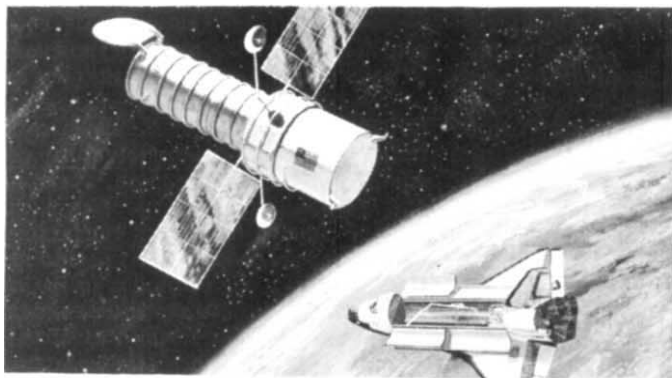
The space telescope. This, as ESA's latest report to COSPAR says, "is the most ambitious project in space astronomy currently planned and will dominate astronomical research for the rest of the century". ESA is participating at a level that will allow Europe's astronomers just 15% of the observing time on the high-resolution, 2.4 metre telescope. The telescope should be launched from the space shuttle at the end of 1983 with five separate focal plane instruments. In a 500-km high orbit, it will be retrievable and refurbishable, and should serve astronomers for two decades.

The future . . .

THE European Space Research and Technology Centre (ESTEC) in the Netherlands is investigating the scientific and technical merit of several spacecraft which will compete for selection as the scientific project to follow the ESA/NASA space telescope which is due for launch in 1983.

To effectively use the science budget ESA's Science Policy Committee (SPC) will have to decide on which experiment to launch by the beginning of 1980. The next mission should then be ready for launch in 1985.

The options are already known. According to Dr D. E. Page, head of the science department at ESTEC, any new mission proposals could not be studied in time for next year's decision



Already approved: space telescope, to be launched from the NASA shuttle, and the Out-of-Ecliptic mission, both scheduled for 1983.

because available staff at ESTEC are already committed. The selection will, therefore, be from the following:

Extreme ultraviolet and X-ray survey satellite (EXUV). The proposal for this satellite stemmed from observations made in 1975 by NASA's X-ray satellite, Copernicus. Before that it had been thought that extreme UV radiation was too heavily absorbed by the interstellar medium (ISM) to be observed. Copernicus, however, showed that the local ISM is thinner than had been expected and that in certain directions the emissions can be observed from objects up to 100 parsec away.

ESA's EXUV satellite would survey the sky in the waveband 15 Å–1,000 Å with a few arc min resolution. It would provide information on very hot stars, which emit most of their radiation in the EUV waveband (250 Å–1,000 Å), and stellar coronae. It would also help study the origins of the diffuse XUV (15 Å–250 Å) background.

The estimated cost on completion of an EXUV satellite is 75 mau. This includes half an Ariane launch.

Astrometry satellite. The position of stellar objects can be measured with far greater precision from space than from Earth. In principle, positional accuracy could be improved by a factor of 20 and the number of observable stars could rise by a factor of 10 to 100.

The idea for an astrometry satellite in ESA is about nine years old. However, it has never been one of the candidates for final selection because previous studies have been abandoned at the mission definition stage. One of the problems is that it could be technically very demanding to achieve the claimed increases in accuracy. Because of this, an astrometry satellite could be expensive, of the order of 95 mau.

Grazing-incidence solar telescope (GRIST). This would be an ESA-funded Spacelab facility to observe the Sun in the soft X-ray and UV wavebands (100 Å–1,700 Å). The advantage of using Spacelab (rather than flying GRIST separately) is that the telescope could be flown as one of several solar telescopes, making simultaneous obser-

vations in a variety of wavelengths. NASA is considering building the other telescopes.

The option of another large Spacelab facility, the large infrared telescope (LIRTS), has recently been shelved on cost grounds. Interest in focal plane instrumentation for LIRTS, however, is still being maintained at ESTEC in case the project is ever reconsidered as a free-flyer. The major disadvantage of putting a heavy instrument on Spacelab is that the launch cost for a 7–30 day flight is as much as the cost of launching it as a free-flyer with an observing time of several years.

Sun-Earth observatory and climatology satellite (SEOCs). This would measure the Earth's radiation budget more accurately than previous NASA satellites. It would help investigate the physical processes controlling the budget by monitoring spatial and spectral changes in the incoming and outgoing radiation. In particular it would monitor the incoming solar UV flux in conjunction with the composition of the stratosphere and observe interactions between different atmospheric levels and the Earth's surface.

SEOCs would be put into circular orbit at a height of 1,150 km and 57.5° inclination. It would only be able to improve on existing measurements if it were to interact with geostationary satellites, such as Meteostat, to obtain information on the diurnal characteristics of the atmosphere. It would also have to interact with polar orbiting spacecraft if it were to study the complete radiation budget. Because the usefulness of SEOCs would depend so heavily on it interacting with other spacecraft, ESA is looking at the possibility of building two.

The cost on completion of one satellite plus the cost of half an Ariane launch is estimated to be at least 85 mau. However, ESA has not yet decided whether to fund climatology out of the mandatory science budget or as an optional applications programme. So SEOCs could be removed from the list of candidates for the next scientific mission and be funded separately.

Polar orbiting lunar observatory (POLO). In spite of NASA's Apollo and Lunar missions, there are still fundamental questions about the Moon left unanswered: for example does it have a core, what is its internal temperature, what is its thermal history, and what was the origin of the magnetic field of the early Moon? Some scientists feel these questions would be better answered with the aid of an orbiting spacecraft than a lander, such as Apollo, which is necessarily very limited in the area it can observe.

A study is currently being rushed through at ESTEC to look at the possibility of building POLO very cheaply from Geos 3 hardware and launching it on the fourth Ariane test flight at the end of 1980. This would mean that POLO could be built cheaply and quickly. It might even mean that it could be funded in addition to one of the other options. The possibility of this happening, however, has grown more remote as cost estimates have risen from a few mau to well above 10 mau.

The SAC will not have to decide which project should follow the one to be approved at the end of next year until the end of 1980. ESA, however, will be sending out a call for mission proposals for that decision in the middle of next month. No one yet knows what that will bring but there is growing pressure from planetary scientists in Europe for ESA to launch spacecraft out of Earth orbit to other parts of the solar system. And POLO is not the only idea they have for the future. They are also considering a cometary mission and a mission to the asteroids. At a workshop organised by ESA earlier this year, European scientists came out in favour of a dual mission to comets Halley and Tempel 2. The proposal they agreed on was to fly-by Halley's comet during its next nearest approach to Earth in 1986 and then, following an Earth swing-by, to rendezvous with Tempel 2 in 1988. Both the proposed cometary and asteroid missions would be extremely expensive. The cost for the cometary mission might be about 500 mau. □

Astronomy in the 1980s

Astrophysics and geophysics are ESA's two oldest disciplines. Professor A. P. Willmore, of the University of Birmingham's Space Research Department presents the case for launching sophisticated equipment to do high energy astronomy . . .

EVEN more than for most sciences, the last two decades in astronomy have been a period of rapid technological development for all its branches. The substitution of high quantum-efficiency detectors for the photographic plate in optical telescopes, the application of large aperture synthesis telescopes in radio astronomy and the entire histories of infra-red, ultra-violet, X-ray and γ -ray astronomy are contained within the last 20 years. The application of space-flight to astronomy, on which the possibility of the newest branches of astronomy (only excluding work in the infra-red) has depended and which promises some important advances in the older branches too, has also occurred during this time.

The last two decades have also been a quite exceptional period of discovery whose many results now demand exploitation. It is also most unlikely that the phases of technological development and discovery are yet at an end. Thus charting a course for astronomy, especially space astronomy which is so costly, will require careful selection from a broad spectrum of projects, any of which in less exciting times might have seemed suitable to be pursued.

Space astronomy—for the obvious reason again of its expense—is highly internationalised; most experiments or facilities involving international teams with members from both sides of the Atlantic. The degree of internationalisation exceeds that even in nuclear physics so that a distinctively European programme is hard to conceive and it is useful only on a world scale to identify the promising paths to develop. But several branches of astronomy are alive and well in Europe, particularly in the UK, and there is no reason to doubt that UK astronomers will be well represented in world class astronomy, provided of course that funding at a minimum necessary level is available.

The sensitivity of large optical telescopes has been substantially enhanced in recent years by the development of photo-electric detectors whose quantum efficiency is far higher than that of photographic plates and whose performance in other respects is comparable. However, the ability of a telescope to detect and resolve objects which are very distant and therefore faint and of small angular diameter depends also on its effective image quality and on the brightness of the night sky. The angular resolution of any good, large optical telescope is determined not by diffraction or the quality of its construction but by the atmosphere, and consequently the Space Telescope, a 2.4m. orbiting telescope being built by NASA with ESA as a minor partner, will have an effective resolution 10 times and a sensitivity 100 times that of the large ground-based telescopes.

The Space Telescope can operate over a wide wavelength range, and will be an ideal instrument for the study of the most distant galaxies which are of course those which are observed in the earliest epochs in the Universe, closest to the time of the Big Bang, and around the epoch of galaxy formation. The measurements of the motions and parallaxes (the apparent motions due to the Earth's orbital motion about the Sun, on which ultimately all astronomical distance scales depend) and of astronomical objects are severely limited by atmospheric effects, and either with the Space Telescope or with satellites constructed for the purpose, improvements of about an order of magnitude should be attainable.

The short history of X-ray astronomy is especially remark-

able for the profusion of new discoveries, for example the close binary sources containing compact stars such as neutron stars and—one supposes—black holes, which allow properties—the mass, at least—of the compact star to be determined in a number of cases. It is salutary to recognise the relative crudity of the techniques which have generally been employed in making these observations, which are about equivalent to carrying out optical astronomy with something between a photographic light meter and a box camera. But much technique development has been carried out in the last 10 years, and the new techniques are now ready to be exploited. Later this year the NASA HEAO-B satellite should be launched with a medium size grazing incidence telescope. This will have an angular resolution roughly two orders higher than any previous X-ray instrument, comparable for the first time with that of an optical telescope. The sensitivity too will be far higher than that of any previous instrument—indeed it will be much better than that of any present X-ray sky survey, which will undoubtedly prove some limitation on its use. This telescope is likely to be able to detect the X-ray emission from clusters of galaxies at distances approaching those of the most distant quasars, thus providing another means of observing the early Universe. Other means of imaging at higher X-ray energies have also been developed, and this is a whole field that is sure to prove productive.

Recent observations of X-ray spectra, which like the spectra of radio sources are dominated by continuum emission, have shown the presence of measurable line emission in a wide variety of sources. This generally arises from highly ionised atoms, usually of Fe, which are present in the source, which must be at a high temperature since it is an X-ray emitter. A high energy feature, however, is believed to be cyclotron emission from around a neutron star, thus allowing for the first time a direct measurement of the large magnetic field (about 10^8 T) around these objects. The diagnostic value of such detailed spectrum observations is very high and together with much improved observations of the spatial structure of sources this is bound to prove an area of important development.

X-ray observations are generally considered to extend up to the electron-positron annihilation radiation at 512 keV. Observations of the γ -ray sky at higher energies are relatively scanty, depending on the surveys carried out with spark chambers on the SAS-2 and Cos-B satellites. Whilst X-ray and optical objects such as the Crab nebula and the Vela pulsar are detected at γ -ray energies, there is a good chance that the typical γ -ray source will prove to be qualitatively different from those observed at lower energies and promising new observational techniques such as drift chambers and double Compton scatter telescopes are now in existence.

In selecting the projects of the future, the obvious criteria to use (and one implicitly accepted in the examples already cited) will be the extent to which they are likely to contribute to fundamental topics—either cosmology, or fundamental physics as general relativity or the properties of nuclear matter. But throughout high energy astronomy, better surveys will be vitally important in the next decade, and will surely bring to light many new types of astronomical object. It will be worth remembering that even at lower energies, both pulsars and quasars were discovered within the last 20 years or so. □

Science in the solar system

... and Dr Alan Johnstone, of the Mullard Space Science Laboratories, poses some of still-unanswered questions about the solar system.

WITHIN the consultative structure of both the European Space Agency (ESA) and the UK Science Research Council, space science is divided between committees at the boundary of the solar system. Although this is a neat arrangement, solar-system science encompasses four almost completely distinct scientific fields, requiring different observational techniques and theoretical approaches. In this article it is not possible to do more than give some idea of the vigour of the research and indicate the principal directions it is taking.

Space plasma physics is concerned with the interaction between charged particles and magnetic fields. The solar wind plasma flowing outward from the Sun forms magnetospheres around the magnetised planets. The magnetospheres contain populations of high temperature plasma confined in the magnetic field.

The particle distributions and the magnetic field configuration of the Earth's magnetosphere are now well-known, but they are not enough to explain the first order operation of the magnetosphere as a plasma machine. The solar wind not only forms the magnetosphere; it is also its principal energy source. The energy is accumulated in the magnetic configuration of the magnetosphere and then released in an explosive event called the magnetospheric substorm, whose most obvious manifestation is the polar aurora. There are many unanswered questions about this process. How is the energy of the solar wind coupled into the magnetosphere and what controls the rate? How is the energy stored and what triggers its release?

The interpretation of satellite observations in such a dynamic system is hampered by the ambiguity between time variations of the whole system and spatial variations along the satellite orbit. The problem can be overcome by deploying a number of satellites in carefully coordinated orbits. This approach has been taken by the ISEE project which has one satellite stationed permanently in the solar wind and a pair of satellites moving close together through the outer magnetosphere to investigate the interaction between the solar wind and the magnetosphere. With the GEOS satellite in geostationary orbit they form a powerful combination to study magnetospheric dynamics.

Another fruitful approach is the comparison of solar system magnetospheres. Each one is dominated by a different effect: Jupiter by its rapid rotation; Mercury by the greater solar wind strength near the sun; and Mars by its atmosphere. The Earth's magnetosphere contains the elements in a dynamic balance. There will be an increase in the use of the plasma in the magnetosphere for plasma physics experiments, especially from Spacelab. The advantages include stable, fully ionised plasma, with no walls to influence the experiment.

The second field is the study of planetary atmospheres. There are now routine observations of the Earth's lower atmosphere for meteorological purposes; and there have been continuing investigations in the upper atmosphere in the regions accessible to satellites. But there have been relatively few observations of the stratosphere and mesosphere in between. A major thrust of atmospheric research is now being directed at the atmosphere between 10 km and 100 km through the international Middle Atmosphere Programme. Our ignorance of this part of the atmosphere became apparent when questions were raised about the influence of various pollutants—such as Freon from aerosol

cans and the exhaust of high-flying aircraft—on the ozone layer.

The chemical reactions which control the composition of the middle atmosphere are extremely complex. Much more information is required about the spatial distribution of various chemical species, the temperature structure, winds, waves and tidal motions before the controlling processes can be understood. All the available platforms will be used in this essentially exploratory work: aircraft, balloons and sounding rockets for in-situ measurements; and free-flying satellites and Spacelab to make global surveys by remote-sensing techniques.

As with magnetospheres, investigating the atmospheres of the other planets gives insight into the origin and evolution of the Earth's atmosphere as well. The two Pioneer Venus space probes will soon be returning detailed information about the dense and hostile Venusian atmosphere.

Solar physicists have made substantial use of space vehicles to study radiation from the Sun at wavelengths from UV to X-ray which do not penetrate through the atmosphere. Their next major project is NASA's Solar Maximum Mission, with 2 European instruments, which is designed to study solar flares during the next maximum of sunspot activity.

The scientific problems are similar to those of magnetospheric substorms. Energy is stored in the magnetic field of Sunspots and then suddenly released in the acceleration of charged particles. Similar questions need answering: how is the flare triggered, and how is the energy transferred to the particles? Since both flares and substorms seem to belong to the same class of plasma processes, any advance in one could assist the other and could also be relevant to other astrophysical and laboratory plasmas.

A second mission, the joint ESA/NASA Solar Polar Probe is going to look at the Sun from a different angle and study the three-dimensional structure of the solar wind. Two spacecraft will use the gravitational field of Jupiter to swing into heliocentric orbits normal to the ecliptic plane taking them over the north and south poles of the Sun respectively.

Finally, the field which has given some of the most spectacular results is the exploration of the planets themselves. US spacecraft are now en-route to Venus, Jupiter, Saturn and Uranus. So far the only European involvement has been by individual scientists but there are projects being discussed now which will be open for possible European participation.

I began by suggesting that the four disciplines within solar system science were grouped together just for administrative convenience. It is true that so far there has been little interaction from them. That may be because the links which must exist have not been fully recognised, but there is an extremely important unifying theme. Each field is concerned with one aspect of the physical environment which makes life possible here on Earth. Why has life developed here and apparently not on the other planets? Is there an important balance which we could unwittingly upset? The tremendous progress that has been made towards understanding all aspects of the origin and evolution of the solar system in the twenty years of research in space science, and the simultaneous realisation that human activities can have damaging effects on a global scale makes it clear that it is a field of research that we ignore at our peril. □

news and views

Insect molecular biology on Crete

from Walter J. Gehring

THIS summer biologists working on the molecular and genetic aspects of insect development converged on the island of Crete for an EMBO workshop held at Kolymbari followed by an international conference organised by the Hellenic Society for Biological Sciences in Heraklion. Crete was chosen because a new University is being established there and also to provide a stimulus to Greek biology. The meetings had a very special ambience which in part was created by the places at which the discussions were being held, which ranged from an ancient church and a Venetian fortress, to a night club and the beach. In the friendly atmosphere generated by our Greek hosts, the open exchange of information proved profitable for all the participants.

From the vast amount of new information presented in Crete only a few highlights can be mentioned.

Structural and functional organisation of the genome

The structural and functional organisation of the genome was the centre of discussion for much of the time. Among the unique genes, the silk fibroin gene has been cloned in bacterial plasmids (Y. Suzuki, Carnegie Institution, Washington) and analysed in detail, especially near the 5' end of the coding region, where one might expect the regulatory signals for transcription and processing of the fibroin messenger RNA. The 5' end of the gene was determined by hybridisation to mRNA which was specifically labelled at the 5' end, in the cap structure. The flanking sequences in front of the gene are AT-rich and contain sequences complementary to the 3' end of 18S ribosomal RNA. Like many other eukaryotic genes, the fibroin gene also contains a large intervening sequence some 1,300 nucleotides long, which is not present in the mature mRNA. Interestingly enough, the DNA sequences at the junction between the coding and intervening sequences are very similar in the mouse β -globin, mouse immunoglobulin gene and the silkworm

fibroin gene, which suggests that the enzymes involved in splicing the mRNA have conserved their specificity in the course of evolution. The notion that fibroin is coded for by a single gene is supported by recent genetic evidence (K. Sprague, Oregon). Genetic variants have been found which differ in the molecular weight of the fibroin. These variants show normal Mendelian inheritance as expected for a single gene, with codominant expression in heterozygotes. Since the protein and presumably also the gene sequence is internally repetitious, these variants in chain length are thought to be due to unequal crossing over.

Several research groups are focusing on the heat shock genes in *Drosophila*. *Drosophila* cells respond to a brief heat treatment by activating a small group of approximately nine gene loci which is reflected in the giant chromosomes by the appearance of specific heat shock puffs. These puffs are sites of intense RNA synthesis and accumulation of RNA polymerase. Under heat shock conditions transcription and translation of the remainder of the genome is reversibly turned off. Pre-existing polysomes disintegrate, a new population of polysomes carrying the newly synthesised heat shock mRNA appears in the cytoplasm, and a set of about eight heat shock polypeptides is synthesised. Obviously, regulation occurs both at the transcriptional and the translational level. This system has considerable potential for studying the coordinate control of a small group of genes. Several heat shock genes have been cloned. In particular, the genes coding for the major heat shock polypeptide of 70,000 daltons have been analysed. Both the analysis of deletions for the major heat shock puffs and evidence from *in situ* hybridisation with cloned DNA indicates that several genes coding for this polypeptide are located at two chromosomal loci, 87A and 87C, where the two major heat shock puffs are located. The restriction maps of the coding sequences at 87A and 87C are similar but not identical. The flanking sequences are very different. Heteroduplex mapping and cross hybridisation studies between the

sequences at 87A and 87C indicate homology of the coding regions and of two small regions apparently outside the coding sequences, near the 5' end of the genes. These sequences are obvious candidates for carrying regulatory signals.

The histone genes of *Drosophila* represent a tandemly repeated gene family (D. Hogness, J. Kemp, Stanford). The repeated element contains five histone genes in the order

H3 H4 H2A H2B H1,
← → ← → ←

where the underlining arrows indicate the direction of transcription of each gene. Such a cluster of five genes is separated from the next cluster by an AT-rich spacer of 1,200 base pairs, the entire element occupying five kilobase pairs. There are approximately 100 repeated elements which, however, do not form a single array, but are instead distributed among several arrays, separated by larger array spacers. Since the histone genes are present in several adjacent bands in the giant chromosomes, the subdivision into arrays may possibly reflect the band-interband organisation. The repeated element containing the five histone genes has been almost completely sequenced. Of particular interest are the short regions between the coding sequences for H2A and H2B, and between H3 and H4. The intergenic sequence between H2A and H2B spans 220 base pairs and contains a sequence of eleven base pairs which is present twice in opposite polarity. This sequence may be a promoter sequence, or part of a promoter.

Another gene family which has been studied extensively is that which codes for the chorion (eggshell) proteins of silkmoths (F. Kafatos, Harvard and Athens). In *Antheraea polyphemus* there are approximately 100 different chorion proteins synthesised and secreted by the follicle cells surrounding the oocyte in a 2-day period during egg development. Each protein is produced during a defined time interval *in vivo* and also in organ culture, indicating that the expression of the chorion structural genes is strictly regu-

Walter J. Gehring is Professor in the Department of Cell Biology at the Biozentrum, University of Basel.

lated according to a developmental program. The proteins can be divided into a few groups A to E. The proteins within each group share homologous amino acid sequences, but differ from each other by amino acid replacements, deletions or insertions. The difficulty of purifying individual mRNAs was overcome by constructing DNA sequences complementary to the mRNA mixture and cloning a pool of hybrid plasmids containing cDNA plasmids for individual messages. Sequencing of some of the cDNA clones revealed previously unsuspected homologies between limited regions of proteins from group A and B. The sequence data suggest that blocks of DNA sequences may be rearranged between genes in the course of evolution. The arrangement of these genes in the chromosomes is not known, but genetic data in the silkworm *Bombyx mori* indicate that the genes for all four classes are located on the second chromosome and probably are clustered (M. R. Goldsmith, University of California, Irvine). The transcription of these genes is being analysed by hybridisation of *in vivo* synthesised RNA to cloned cDNAs using a newly developed spotting procedure.

Transposable genes

Among the most intriguing genes are the dispersed repetitive gene families such as *copia* and *412* in *Drosophila* (D. Hogness, M. Young, Stanford). The *412* and *copia* elements are approximately 7.2 and 5.2 kb in length, respectively, and are found at 20 to 40 different chromosomal sites. The adjacent sequences are different at the different chromosomal locations. These sequence elements are homologous to abundant poly (A)-containing cytoplasmic RNAs. A curious feature of both families is that the elements are flanked by direct repeats of 0.5 and 0.3 kb respectively. The function of these flanking repeats is not known but by analogy with bacterial systems, one can speculate that these elements are translocated to different positions in the chromosomes by reciprocal recombination events which involve the repeated termini. This hypothesis can be tested in a case in which a *copia* element occurs within one unit of a tandemly repeated sequence, by comparing the sequence of the unit with and without the *copia* element. Preliminary evidence obtained by restriction mapping is consistent with the model. Evidence for transpositions of these elements at least on an evolutionary time scale comes from the observation that two wild-type strains differ with respect to the chromosomal location of these elements.

From genetic experiments, 'jumping genes' have been known for some time

RNA processing

from Andrew Travers

THE excision of intervening nucleotide sequences during the production of eukaryotic mRNA and tRNA molecules has recently highlighted the biological importance of post-transcriptional processing. This excision together with postulated subsequent ligation often involves the removal of an RNA sequence at least several hundred nucleotides long.

How do the processing enzymes achieve the exquisite accuracy that is a prerequisite for the maintenance of the continued function of the mature RNA molecules? One possible prokaryotic solution to this problem has recently emerged from the studies of Young and Steitz on the production of the 16S rRNA (*Proc. natn. Acad. Sci. U.S.A.* **75**, 3593; 1978). This rRNA is cotranscribed with the sequences for the 23S and 5S rRNA and is generated by RNase III cleavage of the precursor followed by secondary trimming. The RNase III cleavage sites are contained in the sequences flanking the 16S rRNA gene. By determining the DNA sequences of these regions Young and Steitz have discovered that the proximal and distal sequences are potentially capable of hybridising with each other to form a stem consisting of 26 base pairs. The individual strands of this stem are of course separated by a loop of about 1,700 nucleotides containing the entire 16S rRNA gene. The biological importance of such a structure is given credence by the observation that both proximal, and distal RNase III cuts are contained within the base-paired region. Further, RNase III may require a double helical region of at least 9 base pairs both 5' and 3' to the point of scission, a structure which is lacking in the immediate neighbourhoods of the scission sites. Thus the most probable interpretation is that a site for RNA processing can be generated by the interaction of distantly spaced complementary RNA sequences. The authors suggest that an analogous process may operate during the production of

eukaryotic messenger RNAs.

Cleavage by RNase III does not by itself generate the normal 16S rRNA sequence, yielding only a 17S precursor. For production of the mature 16S RNA two or more additional processing enzymes are required. One recent addition to the family is the endonuclease RNase M16 which cleaves the 16S precursor immediately proximal to the normal 5' end of the 16S RNA molecule (Dahlberg *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 3598; 1978), while yet another enzyme is apparently required for the production of the mature 3' end (Hayes & Vasseur, *Eur. J. Biochem.* **61**, 443; 1976).

A further quirk in the enzymology of RNA processing is provided by the discovery that RNase P, first detected as the activity required for the production of *su⁺III*tRNA (Altman & Smith *Nature new Biol.* **233**, 35, 1971) contains a specific RNA moiety 350 nucleotides long in addition to the expected protein (Stark *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 3717; 1978). As normally purified this highly unusual enzyme contains about 80% RNA by mass. Furthermore the RNA component seems to be essential for nuclease activity. This requirement for the RNA molecule could be trivial. For example, it might be needed to stabilise an active conformation of the protein component. Alternatively the authors suggest the biologically more interesting possibility that the RNA molecule may itself participate in the recognition of the RNase P cleavage site by a mechanism involving direct RNA-RNA interaction. Implicit in such a model is of course the potential for varying the substrate specificity of the enzyme by simply changing the sequence of the RNA moiety. Again, if this enzyme is of universal occurrence such a capability might have important implications in the production of eukaryotic RNA species.

Andrew Travers is at the MRC Laboratory of Molecular Biology, Cambridge.

in *Drosophila*. For example, several transposable elements carrying the *white* locus have been described. *White* was the first mutant to be found in *Drosophila* and 'traditionally' it maps on the X chromosome. However, fly stocks have been found in which the *white* locus is transposed to different chromosomal sites and transpositions are observed in these laboratory stocks. At the meetings, the isolation of a

hybrid plasmid containing complementary sequences to such a transposable element was reported (W. Gehring, Biozentrum, Basel). The element carries both the *white-apricot* allele and the neighbouring *roughest⁺* gene. This plasmid is obviously of great interest for the study of transposable elements in *Drosophila*. It may also provide a general tool for the isolation of genes with no known biochemical function.

Since the transposable element causes mutations by inserting into genes, the flanking gene sequences can be isolated by using the hybrid plasmid complementary to the transposable element as a probe. Alternative methods for gene isolation in general were discussed, in particular the possibility of starting from a DNA sequence which maps close to the gene to be isolated and 'marching' along the chromosome by isolating overlapping DNA sequences (S. Artavanis-Tsakonas & Hogness, Stanford). The present rate of 15 kb per month may be accelerated by using vectors which can accommodate very large inserts of donor DNA.

Polytene chromosomes

The structure and function of polytene chromosomes remains a controversial field to which the recent studies, using fluorescent antibodies to localise RNA polymerases and other chromosomal proteins, have added a new dimension. Classical genetic studies indicate a strong correlation between genes and bands, but the analysis is more or less confined to unique essential genes and it is difficult to estimate what fraction of the genes escapes detection in these genetic tests. In contrast to the hypothesis that the boundary between bands and interbands in the giant chromosomes is defined by specific DNA sequences, a more dynamic model has been proposed on the basis of fluorescent antibody studies (E. Bautz, Heidelberg). This model assumes that puffs and interbands represent the actively transcribed fraction of the genome which varies with the stage of development and between tissues. The well-founded hypothesis that puffs are activated genes has obtained further support by the elegant demonstration of active units of transcription in Balbiani Ring 2 of *Chironomus* by the Miller spreading technique and electron microscopy (B. Daneholt, Karolinska Institute, Stockholm). It appears that there is only one active DNA sequence of 37 kb being transcribed in a chromomere which, on the basis of microspectrophotometry, contains as much as 470 kb.

Endocrinology

Insect endocrinology has also received a new impetus. A plant ecdysone, ponasterone A, has been found with about 50 times the activity of the insect hormone β -ecdysone. Its biological action seems to be identical to that of the natural moulting hormone. Because ponasterone A can be labelled to high specific activity by catalytic tritiation of an unsaturated precursor, stachysterone C, it can be used to identify putative ecdysone receptors, which so far has not been possible (M.-A. Yund, Berkeley). Data from

several laboratories indicate binding constants of the order of 10^8 to 10^9 M. A similar binding constant has been reported for a juvenile hormone receptor. Moulting hormones have now been discovered in the ovaries of several insects and intermediates in the biosynthesis of ecdysone are being isolated and characterised (J. Hoffmann, Strasbourg). The peaks of ecdysone concentration coincide with the time of secretion of the egg shell during oogenesis, and with the time of cuticle formation during embryogenesis. Vitellin, the major yolk protein, seems to be under hormonal control by juvenile hormone in most insects with the exception of mosquitoes and possibly other diptera, where ecdysone seems to take over this role. In locusts, vitellin is a glycolipoprotein of 550,000 daltons which is derived from a precursor, vitellogenin, that is synthesised in the fat body (G. Wyatt, Kingston, Ontario). *In vitro* studies indicate that initially two polypeptides of 265,000 and 250,000 daltons are synthesised which may represent products from two different genes. The labelled precursor can be chased progressively into a group of smaller polypeptides of 140,000 to 52,000 daltons. Vitellogenin synthesis is induced by juvenile hormone, secreted by the fat body and transported to the oocytes, which makes the system analogous to vitellogenin synthesis in vertebrates. Since methods are now available for the isolation of hormone-inducible genes and several ecdysone-inducible genes have already been isolated in *Drosophila* (Kemp & Hogness, Stanford), one can expect rapid progress in this area.

In the area of insect cell culture there has also been considerable progress. *Drosophila* cell lines have now been established from several mutant strains (H. P. Bernhard, Biozentrum Basel), and successful cell fusion has been announced (C. Wyss, ETH Zurich). Perhaps the most important recent contribution has been the establishment of a mutant cell line with the haploid chromosome number (G. Echallier, Paris).

Although there has been an extensive search for regulatory mutants, *Drosophila* geneticists have so far failed to provide conclusive evidence for the existence of regulatory elements controlling the activity of structural genes in development. It may be that our ideas have been too greatly influenced by models developed for prokaryotic systems. Interesting new results on genes controlling compartment formation in imaginal disks (T. Kornberg, University of California, Los Angeles; P. Lawrence, MRC Laboratory of Molecular Biology, Cambridge) and mutants interfering with sex de-

termination were presented (R. Nöthiger, Zürich). If such genes can be isolated by molecular cloning techniques, it may become possible in the future to unravel controlling mechanisms operating at these higher levels.

The workshop and the conference were among the best meetings that I have attended, especially in promoting the dialogue between scientists working in related fields, and our Greek hosts merit special thanks for providing an ideal setting which made these meetings so pleasant. □

Ice dynamics

from a Correspondent

Is man's effect on world climate likely to precipitate a catastrophic rise in world sea level as suggested by Mercer (*Nature* 271, 32; 1978) and is a major surge of the Antarctic ice sheet still considered possible by glaciologists? These and other questions were discussed during the recent International Glaciological Society's Symposium* on 'The dynamics of large ice masses' at Carleton University, Ottawa. Thanks to the associated 'Symposium on glacier beds: the ice-rock interface' organised under the National Research Council of Canada at the same location during the preceding week, there were many healthy exchanges between glacial geologists who had studied the effect of ice dynamics in the past, glaciologists studying present day glaciers and ice sheets and an increasing number of glaciological modellers. Most participants benefitted from attending both symposia, a happy juxtaposition of meetings since the glacier bed is the inadequately understood base of glacier dynamics in more ways than one.

It is apparent that while our understanding of the internal deformation of ice sheets is advancing along satisfactory lines, our knowledge of how much sliding takes place at the glacier bed, how much it is controlled by rock debris in basal ice, by basal water pressures and by pressure melting and regelation are still far from adequate. Until this is cleared up, modellers will lack guidance on how to treat the base of the ice sheets of Greenland and Antarctica. In this connection, long awaited reports on the debris-laden ice cores from the base of the ice sheets at Camp Century, Greenland and Byrd Station, Antarctica were of special interest.

A. J. Gow and W. Sheeby (US Army's Cold Region Research and Engineering Laboratory, Hanover, New Hampshire) and S. Epstein (Caltech) showed that the lowest 4.83 m of the core from Byrd Station contrasted with

* Held 21-25 August, 1978.

ice immediately above through its very low content of dissolved air, the presence of rock particles ranging from clay size through sand to pebbles and cobbles, and in its oxygen isotopic composition. The latter was around 4‰ lower in $\delta^{18}\text{O}$ values in the basal ice, and more variable than in the ice above. Thus when one allows for fractionation effects it seems that the basal ice that was formed by refreezing of the melt water was of fairly local origin, but it is not clear whether freezing took place in a single period or was episodic. Melt water was present beneath the ice when the base was reached in 1968.

Whereas debris weight percentages of up to 15.7% were present in the basal ice from Byrd Station, Susan Heron and C. C. Langway (Buffalo University) found an average debris concentration of only 0.24% in the bottom 15.7 m of the 1966 core from Camp Century. This contained over 300 alternating bands of clear and debris-laden ice. Although the basal ice temperature is now -13°C , they conclude that the debris was incorporated by freezing of subglacial water over a basal till. There was only one pebble in the Camp Century ice, the remaining fine debris being either distributed throughout the dirty ice layers or included in the form of pebble-sized aggregates of fine particles picked up from the basal till. In contrast to the Byrd core, the gas concentration in the basal ice at Camp Century was about half that at higher levels. The gas showed traces of methane, high CO_2 levels, low O_2 levels and enrichment of argon. It was suggested that although air may have been excluded during freezing, its composition indicated subsequent diffusion into the basal ice.

As with glacier beds, the main advances in our understanding of the dynamics of the Antarctic ice sheet are coming from field measurements often involving new techniques. Radio echo mapping from long range aircraft, satellite altimetry and Doppler position fixing as well as more conventional survey methods are providing much better knowledge of the form and flow of the polar ice sheets than seemed possible a few years ago. N. Young (Melbourne University, Australia) who took part in Soviet traverses over the east Antarctic ice sheet during the past two summers found that ice velocities measured by satellite Doppler techniques agreed with calculated balance velocities within an accuracy of 20% in a region of around one million km^2 between Mirny Station and Dome 'C'. Other reports indicate that areas feeding the Lambert and Shiraze Glaciers may not be in balance, while

evidence discussed by K. Rose of the Scott Polar Research Institute suggested that an ice stream feeding the Ross Ice Shelf from Byrd Land may have surged in the not too distant past. The general consensus appeared to be that while surges of individual glacier basins in Antarctica might occur, the idea of a single massive surge of 10 to 20% of the total ice mass of Antarctica as a cause of ice ages (Wilson *Nature* **201**, 147; 1964) was losing ground.

Laboratory studies on ice cores from Law Dome by D. S. Russell-Head and W. F. Budd (Antarctic Division, Department of Science, Melbourne) showed that flow (shear strain) rates in ice with strongly aligned C axes were typically three to five times faster than in randomly-oriented polycrystalline ice, whereas crystal size had little effect on flow rates. These results matched well with measurements of borehole inclination from the same locality. It is becoming clear that the variation of horizontal ice velocity with depth is more complex than expected and that the accuracy of age-depth relationships in ice cores based on ice flow models is limited by these factors.

The above did not, however, deter the modellers who use computer simulations to increase our understanding. Of these, a catastrophe model of the palaeoclimate by D. R. MacAyeal (University of Maine, Orono) that considered changes of insolation, sea level and isostatic depression of the glacier bed helped to explain the sudden transitions between ice-free and ice age conditions.

In addition to ice sheets, floating ice shelves and sea ice were discussed in half day sessions. Towed side scan sonar records from a ship obtained by O. Orheim of Norsk Polarinstitutt showed both basal scour of the sea bed by icebergs and a fascinating pattern of irregular melting at a depth of 100 to 200 m beneath straight cliffs of the ice front. Discussions of the mechanical and thermal stability of ice shelves and marine ice sheets indicate that before any sudden climatic warming would release the ice sheet of West Antarctica and affect sea level there would be a delay of some hundreds of years. After this, considerable progress in modelling of Arctic pack ice seemed relatively mundane, if no less important climatologically.

Overall, it is becoming clear that to advance from a few elementary but over-simplified concepts of glacier behaviour to a detailed understanding of the dynamics of ice sheets involves more complex investigations than was commonly realised a decade ago, but much progress is being made. This is significant since hydroelectric projects

and the development of Arctic oil involves judgement on glaciological matters that are of major importance to industry. □

New magnetic structures in the light rare earth metals

from S. B. Palmer

RECENT weeks have seen two significant developments that have cast new light on the magnetic behaviour of the rare earth elements, praseodymium and neodymium. At room temperature both elements are paramagnetic with localised magnetic moments produced by the electrons in the partially filled 4f shell. In the majority of the rare earth metals the exchange interaction is sufficiently strong to produce magnetic order at low temperatures although the oscillatory, long range nature of the exchange interactions results, in many cases, in complicated modulated magnetic structures. Each magnetic ion also experiences an electrostatic field with the symmetry of the crystal lattice: this crystal field interaction opposes the exchange by favouring a uniform arrangement of the moments. For both neodymium and praseodymium the crystal field and exchange energies are comparable, but in neodymium the exchange energy is sufficiently strong to produce magnetic ordering below 20 K.

The magnetic structure of neodymium has been a mystery since the initial neutron diffraction measurements of Moon *et al.* (*J. appl. Phys.* **33**, 1041; 1964) carried out at Oak Ridge. The double hexagonal close packed (dhcp) structure has two separate sublattices, one with cubic and the other with hexagonal symmetry. Moon proposed that at 20 K the ions on the hexagonal sites ordered antiferromagnetically. These sites form planes lying perpendicular to the hexagonal axis and in a particular basal plane the magnetic moments, according to Moon *et al.*, are confined to the plane and oriented along the crystallographic *b* direction. The moments are sinusoidally modulated with the *q* vector of the modulation in an *a* direction, at 90° to the moments.

This model, however, was inadequate to explain all the features of both Moon *et al.*'s results and the more recent work of Lebech and Rainford (*J. Physique* **32**, 370; 1971; *Proc. Int. Conf. Mag. Moscow* **3**, 191; 1974.) A

S. B. Palmer is in the Department of Applied Physics, University of Hull.

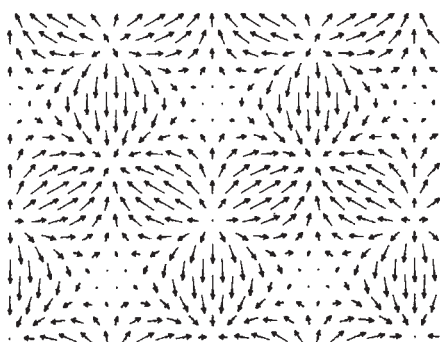


Fig. 1 Magnetic structure of neodymium for $\alpha=0^\circ$. The basal-plane components ($\uparrow$) of the spins in the A layers centred at the atomic positions of the undistorted lattice.

better fit to the experimental data could be obtained if it was assumed that the cooperative order on the hexagonal sites induced magnetic order on the cubic sites although the results were still far from satisfactory.

The report by Bak and Lebech (*Phys. Rev. Lett.* **40**, 800; 1978), of a magnetic structure that fits all the experimental results brings relief to many solid state physicists who have spent long hours attempting to unravel the neodymium puzzle. The novel magnetic structure of the hexagonal sites is shown for one basal plane in Fig. 1 where the magnetic moments are modulated in the three equivalent ' a ' directions simultaneously, hence the name of 'triple- q ' modulated magnetic structure. The work of Bak and Lebech in Denmark involved experimental proof that the magnetic phase change at 20 K was of second order, hence allowing Landau symmetry arguments to be used with renormalisation group theory to predict that the three different equivalent q vectors are present simultaneously rather than contained in separate domains as in the structure of Moon *et al.*

Associated with the magnetic structure shown in Fig. 1 are lattice distortions that are identical in form except for a phase difference of 90° . Initial estimates indicate that these distortions may well be of the order of a few percent and as such should be detectable by X-ray scattering studies.

Measurements of specific heat, elastic constants, magnetoresistance and neutron diffraction have all revealed a multitude of phase changes as a function of applied field in the temperature range 1-20 K. The extension of Bak and Lebech's model to high fields is eagerly awaited.

In contrast to neodymium, praseodymium is fascinating, not for the order it exhibits, but for the magnetic order it nearly exhibits. Experiments on single crystal praseodymium demonstrate that it remains paramagnetic down

to 30 mK: the exchange is about 90% of the critical value for magnetic ordering. In contrast polycrystalline praseodymium, highly strained because of anisotropic thermal contraction, orders antiferromagnetically at 25 K. The addition of 3% neodymium to single crystal praseodymium produces order below 6 K. Because of the large orbital angular momentum of the Pr ions, there is a powerful coupling between the lattice and spin systems. Jensen's analysis (*J. Phys. C* **9**, 111; 1976) of this magnetoelastic coupling led him to predict that, at 4.2 K, a relatively modest pressure of 1,000 bar applied along the crystallographic a direction would, by distorting the basal plane, change the delicate balance between exchange, crystal field and magnetoelastic energies and induce magnetic ordering. In a recent investigation of these predictions McEwen *et al.* (*Phys. Rev. Lett.* **41**, 343; 1978) have carried out (at the Institut Laue-Langevin, Grenoble) a series of careful neutron inelastic scattering studies of praseodymium as a function of uniaxial stress applied along the a axis. They found that as the stress was increased the

lowest energy magnetic excitation dramatically decreased, heralding the antiferromagnetic order which was established at 800 bar. The presence of the ordered phase was confirmed by the observation of elastic satellite reflections corresponding to a long-range magnetic structure. At 800 bar the Néel temperature was 7.4 ± 0.4 K. On removal of the stress the magnetic satellites were absent and the exciton dispersion curve was identical to that of previous workers and to the curve obtained before stressing took place. Both these results indicate that there was no residual stress left in the sample and the process was completely elastic.

The magnetic structure postulated by McEwen *et al.* is a basal plane modulated structure, identical to that of Pr - 3% Nd and reminiscent of the neodymium structure discussed above. Jensen's elegant theories, which have already led to quantitatively accurate explanations of neutron inelastic scattering studies of exciton phonon coupling, field dependence of elastic constants and magnetostriction, seem to have been vindicated again. □

Sun-weather effects

from Roger H. Olson

MOST of the papers presented at a recent conference on Solar-Terrestrial Influences on Weather and Climate* reflected the fact that Sun-weather research is still pretty much in the empirical stage. There is even disagreement on the details of many of the empirical findings. For many years the conventional wisdom has held that Sun-weather effects were stronger in winter than in summer, and stronger in the auroral zone than in low latitudes. However, this may be just a reflection of the fact that changes are more dramatic in high latitudes and in winter, and empirical studies tend to focus on the more dramatic effects. As far back as 1965, H. C. Willett of the Massachusetts Institute of Technology pointed out that, when properly normalised, the weather data show stronger responses to solar signals in summer and in low latitudes. Although there is still conflicting evidence, Willett's ideas received strong support from several investigators. There was also support for the idea that the X-ultraviolet portion of the solar spectrum may be more important than the particle emission, that is, solar proton events.

As for low energy solar particles, the so-called auroral particles, there was relatively little discussion of geomagnetic-weather effects. One possible reason is that in using geomagnetic data there is the danger that a small part of the geomagnetic signal could be caused by a weather disturbance. Although this danger is rather slight, many investigators prefer to use solar signals that are unmistakably solar, such as solar flares and plagues, sunspots, and magnetic sector boundaries. One particular class of geomagnetic disturbances did receive special attention, however—the Forbush decrease events. In a Forbush decrease, the magnetic disturbance causes a decrease in cosmic ray flux throughout the whole depth of the atmosphere, including the surface layers. Thus, in seeking mechanisms to explain the various Sun-weather effects observed, the Forbush decrease is becoming more and more interesting.

One theme that ran through the conference was the importance of Sun-weather effects in weather forecasting. In 1977 M. F. Larsen and P. M. Kelley of Cornell University found that the accuracy of US weather forecasts took a dip one day after solar sector boundary crossings. Following their lead, R. Reiter (Institute for Atmospheric Environmental Research of the Fraunhofer Society, FRG) reported that German forecasts suffered a

* The first Ohio State University Sun-Weather Conference was held on the campus of the University from 24-28 July, 1978. The conference was organised and chaired by T. A. Seliga, Ohio State University, and B. M. McCormac, Lockheed Palo Alto Research Laboratory. It is hoped that this type of conference can be repeated every two to three years.

similar fate. J. W. King (Appleton Laboratory, UK) reported similar findings for British forecasts. All these results cover relatively short time spans and limited geographical areas, and thus should be treated as tentative, but highly suggestive.

There was a healthy degree of skepticism expressed by several speakers. W. O. Roberts (Aspen Institute for Humanistic Studies, Boulder) stressed that, in spite of the virtual certainty that Sun-weather relationships exist, we desperately need rigorous theoretical studies and models to explain them. J. Namias (Scripps Institution of Oceanography, La Jolla) chose to focus on one meteorological parameter, namely drought. He stressed the importance of understanding the full implications of the physical causes of droughts in attempting to correlate drought with solar activity. Interestingly enough, although drought is typically associated with anticyclones, he felt that the best place to look for Sun-weather influences is in cyclonic activity. His attitude is perhaps typical of many highly skilled and experienced meteorologists—namely that Sun-weather research, in spite of its long and frustrating history, deserves careful scrutiny. The chief skeptic by far, was A. B. Pittock, temporarily of the University of Arizona. He cast strong doubts on many published reports of correlations between weather and the various solar cycles (such as 27 days, 11 years and 22 years). The 27-day cycle was also attacked by others, who pointed out that it is difficult to distinguish between 27-day solar and lunar effects. Pittock had little quarrel with some of the short-term correlations which have been reported recently.

As regards mechanisms, the cupboard is not entirely bare. Perhaps the most popular was the thunderstorm mechanism, discussed by R. Markson (Massachusetts Institute of Technology) (see *Nature* 273, 103; 1978) and others. He regards the thunderstorm, and its associated high resistivity to vertical electrical currents as a valve controlling the whole electric circuit of the atmosphere. Solar flares, and possibly Forbush decreases, can modify the resistivity above thunderstorms, leading to profound changes in the global circuit. This could feed back into the physics of the clouds, affecting precipitation patterns and affecting gross features of the atmospheric energy balance. Empirical evidence supporting this idea was furnished by Markson himself, by R. Reiter, and

by M. Lethbridge (Pennsylvania State University). They all presented evidence that thunderstorm frequency is related to solar signals such as flares and solar sector boundaries.

Ozone was also mentioned as a possible trigger mechanism. M. E. Schlesinger (Oregon State University) suggested that an initial ozone increase, caused by solar activity, could lead to circulation changes, which in turn would lead to advective changes in ozone concentration, this leading to a positive feed-back. Many others agreed that some sort of ozone mechanism might be useful. Another possible feed-back mechanism was suggested by E. C. Kindle (Old Dominion University, Norfolk, Virginia) and also elaborated by others. This involves the Charney-Drazin effect, which operates on about a 10-day time scale, a typical time scale for some reported Sun-weather effects. If the speed of the middle latitude westerlies is decreased, this allows greater penetration of tropospheric wave energy up into the stratosphere. Thus if the troposphere is subjected to an influence (possibly a solar influence) that causes it to lose kinetic energy, the feed-back mechanism

will cause it to lose even more energy through dissipation to the stratosphere.

The symposium ended with a number of resolutions. One involved a request to international scientific organisations, such as ICSU, to place the Sun-weather problem high on the list of future projects. Another was to decry the tendency for old, reliable data sets to be threatened with extinction, due to retirement of key personnel or lack of funding. Examples of data sets which are either threatened or actually ended include the Zurich sunspot observations, the Ottawa centimetric radio observations, the Greenwich daily sunspot drawings, and the McMath daily plage observations. These have been of great value both to research workers and to those engaged in operational solar-terrestrial problems. As an example of the latter, there were reports given by two commercial weather forecasters who use solar data in making their forecasts. Although it may be regarded as premature, even by some enthusiasts, to use solar flares, for instance, in routine weather forecasting, the fact that it is being done underscores the need for a serious study of the problem.

Therapeutic agents for sickle cell disease

from M. F. Perutz, J. Rosa and A. Schechter

At a recent workshop in Paris* the molecular and cellular bases of sickle cell disease and possible new approaches to its therapy were discussed.

M. F. Perutz (MRC Laboratory of Molecular Biology, Cambridge) reviewed X-ray diffraction and electron microscope studies of the gels of deoxyhaemoglobin A that precipitate in the red cells of patients with sickle cell anaemia. These gels consist of fibres of 20–22 nm diameter, each made up of several filaments of 6 nm diameter in which single haemoglobin molecules are strung together like pearls. The structure of the contacts between successive molecules in these filaments is known in atomic detail from X-ray analysis of crystalline deoxyhaemoglobin A in which the molecules are strung together in the same way. Sickling is probably caused, first, by the pairwise aggregation of such filaments, as in crystals of deoxyhaemoglobin S (deoxyHb S), where the valines of one filament touch nonpolar side chains between helices E and F β in the neighbouring filament of the pair, and then by the aggregation of several such pairs to form a fibre. Dykes *et al.* proposed

that each fibre is a rope made up of 14 filaments (*Nature* 272, 506; 1978); A. Klug suggested that this number may arise from the winding of six pairs of filaments around one central pair.

W. Love (Baltimore) reported that the resolution of the X-ray analysis of single crystals of deoxyhaemoglobin S had been raised from 5.0 to 3.0 Å, good enough to distinguish individual residues. The new map confirms that the structure of deoxyHb S is the same as that of deoxyHb A, except that the substitution of valine for glutamate 6 β may have forced the first turn of helix A to a conformation different from that in deoxyhaemoglobin A. B. Magdoff-Fairchild (New York) reported her exciting discovery that oriented, paracrystalline preparations of sickled cells or deoxyhaemoglobin A can turn into bundles of crystals isomorphous with those analysed by Love and his colleagues. The intensity distribution in the X-ray diffraction patterns of these crystalline bundles is similar to that of oriented gels, suggesting similar molecular arrangements. This raises hopes that by combining information from crystals and gels we may be able to find out more about the structure of the contacts between molecules in the

Roger H. Olson is at the Aspen Institute for Humanistic Studies, Boulder, Colorado.

* Held on 19–21 July and sponsored by INSERM and the US National Institutes of Health.

gels than has been possible by studies of gels alone.

At constant temperature solutions of deoxyHb S remain liquid for long periods until they suddenly gel. The delay is exactly reproducible and is inversely proportional to high powers of the temperature and haemoglobin concentration (*Proc. natn. Acad. Sci. U.S.A.* **71**, 4864; 1974). W. Eaton (Bethesda, Maryland) proposed to use this phenomenon for testing anti-sickling drugs. To make the disease significantly less severe, a drug would have to increase the delay 10–100-fold, equivalent to a 1.06–1.17-fold increase in the solubility of deoxyHb S or a 1.5–4-fold decrease in intracellular Hb concentration or a 4–7-fold increase in oxygen affinity. These are stiff requirements. E. R. Heuhns and M. Rosemeyer measured the antigelling properties of Hb A, F and A₂ when added to Hb S, and found that the antigelling effect rises in that order. This suggests that activation of the gene for Hb A₂ would have a more potent antisickling effect than activation of the gene for Hb F.

Cell studies

Studies of whole normal and sickled cells were reviewed next. J. Hoffman (New Haven) discussed the distribution of diffusible ions and control of cell volume in normal erythrocytes, and the use of fluorescent dyes to study the intracellular milieu. Such studies are very relevant in view of the steep dependence of the gelling delay on haemoglobin concentration. M. Mohandas and N. Bessis (Paris) used anisotropy of laser light scattering to measure the deformability of red cells sheared between the concentric cylinders of a viscometer. This method could be used to measure the degree by which antisickling agents diminish the stiffening of deoxygenated sickle cells.

The physiology of the sickle erythrocyte is more complex than that of Hb A solutions. Recent technical advances are now allowing their oxygen equilibria to be measured. R. Bookchin (New York) found that the Bohr effect of sickle cells is steeper than that of normal ones and that 2,3-diphosphoglycerate, the intracellular oxygen regulator, increases gelation merely by lowering the oxygen affinity and not, as had been reported, by making deoxyHb A less soluble. R. Benesch and R. E. Benesch (New York) measured the oxygen affinity of Hb A as a function of haemoglobin concentration and found a discontinuity at

the point of gelation corresponding to the minimum gelling concentration. The concentration at which this discontinuity occurs could be used as a measure of the antigelling properties of drugs. If any drug inhibited sickling without raising the oxygen affinity of solutions which are too dilute to gel, then this would prove that it acted on the intermolecular contacts.

Antigelling and antisickling agents

Compounds being tested *in vitro* as antigelling or antisickling agents were discussed next. These may act by more than one mechanism, but can be classified according to their primary effects. The first class consists of modifiers of Hb A which raise the oxygen affinity. The prototype of such reagents was KNCO which proved too toxic for oral therapy. According to L. Kraus carbamylphosphate reacts with haemoglobin in the same way as cyanate, but the unreacted fraction may be less toxic as a result of enzymatic degradation. Imidoesters, both bifunctional dimethyladipimide (B. Lubin, Oakland, California) and monofunctional methylacetimidate (T. Gabuzda, Philadelphia) decrease sickling by raising the oxygen affinity; so does cystamine which reacts with cysteines 93 β (Y. Beuzard & J. Rosa, Paris). Among disulphide reagents cystine dimethyl-ester reacts with the same groups, increases the oxygen affinity and has antisickling properties demonstrable in single erythrocytes (D. L. Currell, Rome). The toxic nitrogen mustards react with histidine 2 β which is one of the diphosphoglycerate binding groups. (E. Roth, New York).

The second class consists of modifiers of intermolecular contacts in deoxyHb S fibres. J. Manning (New York) has found that glyceraldehyde inhibits sickling, probably by reaction with a specific lysine not yet determined. It also acts on the membrane. Other compounds in this class interact non-covalently, such as the alkylureas studied by R. Nagel's group (New York) which are more potent inhibitors than urea itself. A. Schechter (Bethesda) has sought stereospecific amino acids or peptides that inhibit gelation and found that aromatic amino acids or peptides containing aromatic amino acids, but not the sickling peptide (residues 1–16 β), have antigelling properties. C. F. Poyart (Paris) found that methane, ethane, propane and butane increased the oxygen affinity by 10–15%, propylene by 25% and butane by 50%; these agents may also act on the intermolecular contacts. A. Cerami (New York) reported experiments with *n*-carboxyaspirindimethylester which reacts with an as yet undetermined group.

The final class of compounds dis-

cussed were modifiers of the whole erythrocyte. G. Brewer (Ann Arbor) described clinical and laboratory studies of Zn acetate as an antisickling drug and C. Natta (New York) discussed DBA(3,4-dihydro-2,2-dimethyl-2H-1-benzopyran-6-butyric acid).

Many of these compounds are too toxic for oral use and would have to be administered extracorporeally, by removing the red cells from the circulation, reacting them with the drugs, and putting them back again. S. Balcerzak (Columbus) has tested extracorporeal administration of cyanate to erythrocytes of dogs and two baboons and found no toxic effects in periods of up to one year. D. Diederich (Kansas) tried extracorporeal cyanate treatment on about 10 sickle cell patients and found a reduction in crises during treatment, as compared to before and after.

According to W. Rosse's (Durham, North Carolina) review of the American clinical trials of orally administered urea, alkali and cyanate, none of these reduced the severity or frequency of sickle cell crises, but the trials helped to establish reliable criteria for evaluating new therapies. At a final round table conference clinicians from West Africa and the USA dwelt on the great variability of the disease, and the absence of reliable information concerning its mortality, morbidity and clinical manifestations, especially in Africa and the Middle East. They were also concerned about the ethical problems of trying antisickling drugs on patients. They would not try them on children and felt that even adults might be better helped by conservative treatments. □



A hundred years ago

FROM the 16th to the 18th of August the city of Cartago in Costa Rica, in Central America, was visited by five distinct shocks of earthquake. No particular damage was caused. What is to be noticed in the meteorological result: that the weather was changed, being attended by heavy showers that flooded the rivers. In all earthquake countries it is believed the weather is affected by such phenomena, and it is desirable that observations should be recorded, showing the influence of earthquakes on weather. It has been alleged that the Comrie earthquakes in Perthshire came on after rains and floods of the river.

From *Nature* 3 October, 18, 600; 1878.

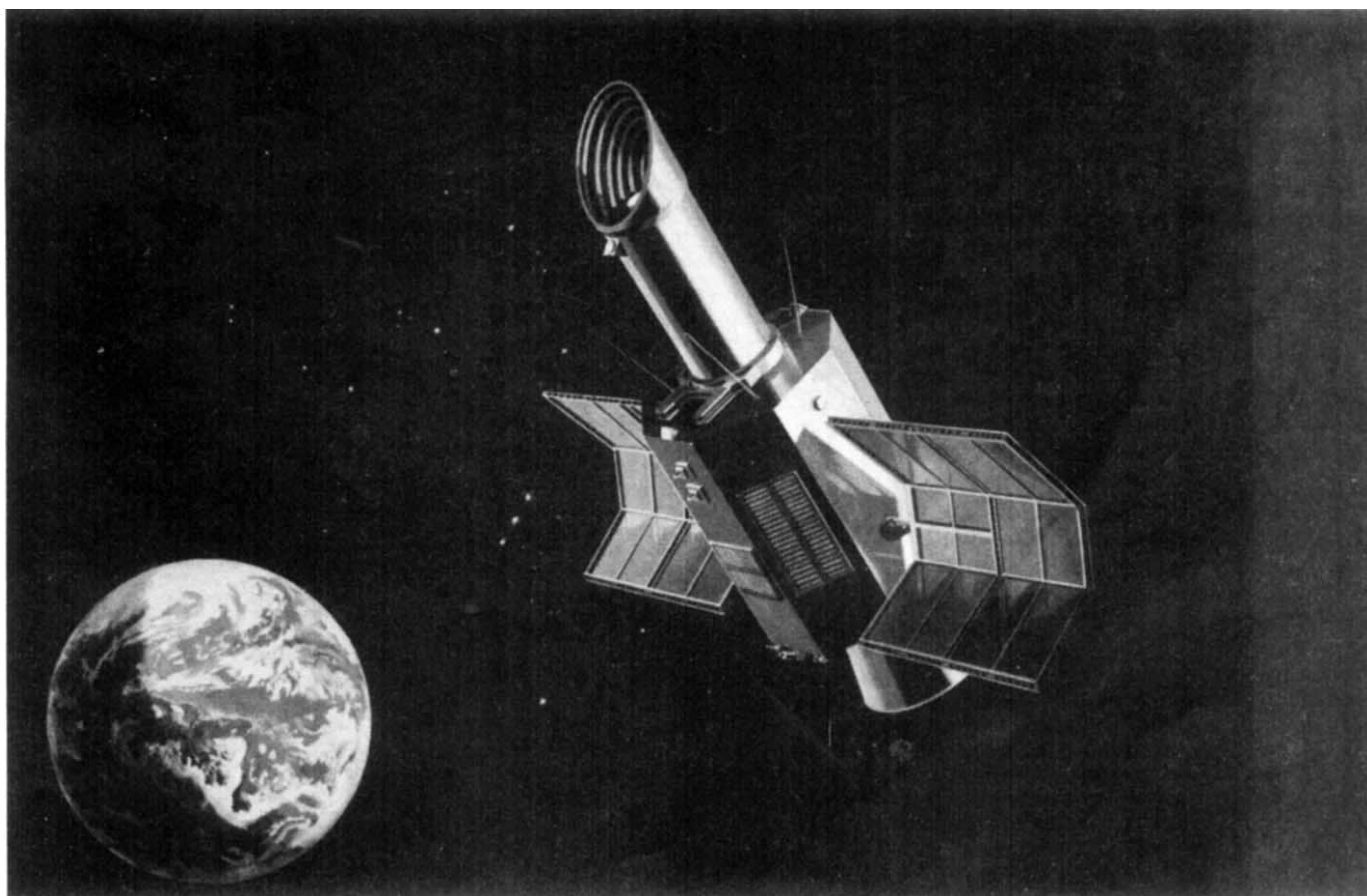
M. F. Perutz is Director, MRC Laboratory of Molecular Biology, Cambridge. A. N. Schechter is at the Laboratory of Chemical Biology, NIAMDD, Bethesda and J. Rosa is Director, Unite et Recherche sur les Anémies, Hôpital Henri Moldur, Creteil.

The International Ultraviolet Explorer (IUE)

The following eight papers deal with the International Ultraviolet Explorer (IUE), a joint project of the National Aeronautics and Space Administration (NASA), the European Space Agency (ESA) and the Science Research Council (SRC). The first paper describes the overall project and concept. The second deals with its performance in orbit as assessed during the commissioning phase following the successful launch of 26 January 1978. In addition to engineering switch-on, sub-system tests, optimisation and calibrations, that commissioning phase included astronomical observations of a set of high-priority targets as an insurance against premature failure of the system. What is 'high priority' is always a matter of judgement, and in this case that judgement was made by an international science commissioning team drawn mainly from the project staff of the three agencies but also including scientists from the general community. The results of those high-priority observations and their subsequent analysis are presented in the final six papers of this series. It should be stressed that these observations were made immediately after switch-on, that is before the system had been fully optimised or calibrated, and that the data were reduced in the very earliest phases of the image processing system. Nevertheless they clearly indicate the potential of IUE and provide a range of new and exciting data in many areas of astronomy.

In addition to the authors listed on the following papers, many people have made essential contributions to this project in all its stages, from conception and approval through hardware production and launch to the recent commissioning and operational phases. To all those people, in the government agencies, research institutes, universities and industry of many countries, the authors extend their thanks and gratitude. These papers were submitted on 9 August 1978.

R. Wilson



The IUE spacecraft and instrumentation

A. Boggess, F. A. Carr, D. C. Evans, D. Fischel, H. R. Freeman, C. F. Fuechsel, D. A. Klinglesmith, V. L. Krueger, G. W. Longanecker, J. V. Moore, E. J. Pyle, F. Rebar, K. O. Sizemore, W. Sparks, A. B. Underhill, H. D. Vitagliano & D. K. West

Goddard Space Flight Center, Greenbelt, Maryland 20771

F. Macchetto & B. Fitton

Astronomy Division, ESTEC, Noordwijk, Holland

P. J. Barker, E. Dunford, P. M. Gondhalekar, J. E. Hall, V. A. W. Harrison, M. B. Oliver, M. C. W. Sandford, P. A. Vaughan & A. K. Ward

SRC Appleton Laboratory, Slough, Berks, UK

B. E. Anderson, A. Boksenberg, C. I. Coleman, M. A. J. Snijders & R. Wilson

University College, Gower St, London WC1, UK

The International Ultraviolet Explorer satellite operates as an astronomical observatory to provide both high and low resolution UV spectra of sources other than the Sun. The spacecraft, its instrumentation and the ground control system are described.

THE International Ultraviolet Explorer (IUE) was launched successfully on 26 January 1978, from Cape Canaveral, Florida. The satellite was developed to provide a general facility for observing ultraviolet (UV) spectra of astronomical sources over the wavelength range from about 1,150 Å to 3,200 Å. The project has been a joint undertaking in which the US National Aeronautics and Space Administration (NASA) provided the spacecraft, the optical and mechanical components of the scientific instrument, the US ground observatory and spacecraft control software; the UK Science Research Council (SRC) in collaboration with University College London (UCL) provided the television cameras used to record the spectroscopic data; and the European Space Agency (ESA) provided the solar arrays and European ground observatory. The image processing software was developed jointly by NASA and the SRC's Appleton Laboratory. The satellite has been placed in geosynchronous orbit over the Atlantic Ocean and is operated for 16 h each day, for NASA sponsored observers, from the US ground observatory located at the Goddard Space Flight Center (GSFC) near Washington, D.C., and for the remaining 8 h by ESA, for ESA and UK sponsored observers, from the European ground observatory located near Madrid.

Here the spacecraft and its instrumentation are described. In the following series of papers IUE's performance in orbit will be discussed¹, and the initial results from early observations taken during the satellite checkout and commissioning period reviewed²⁻⁷.

Design concepts

It was decided that a high spectral resolution of the order of 0.2 Å would be required to study the atmospheric characteristics of bright stars and planets and that a lower resolution of about 6 Å should be available to obtain information on faint sources. In addition, two other requirements were fundamental in determining the design of the scientific instrument and of the spacecraft to support it: first, the instrument should be able to record simultaneous data over large spectral ranges and, second, the spacecraft should be able to take maximum advantage of the observer's real-time judgement concerning the quality and content of his data.

The first suggests a spectrograph using an imaging system, such as a television camera, capable of integrating the spectrum. A spectrograph with echelle optics gives a compact spectral

format with high dispersion and, since the echelle image consists of many adjacent spectral orders displayed in a closely spaced array, it makes efficient use of the field of view of a television camera. To achieve adequate resolution, the total spectrum must be split into two ranges, from 1,150 Å to 1,950 Å and from 1,900 Å to 3,200 Å, and so two cameras are required to record the full spectral range. The observing efficiency is hardly affected by this split, as two exposures are usually required to expose, optimally, both the long and short wavelength spectra.

The second requires good communication between the spacecraft and the observer so a geosynchronous orbit was chosen. The IUE orbit allows continuous real-time control from the NASA control centre and at least 10 h each day from the ESA control centre. The arrangement is fundamentally different from low altitude orbiting observatories which have to acquire data while remote from ground control. In the case of IUE, control and performance monitoring are exercised continuously. The telescope field is displayed to the observer, who identifies his target star and directs the course of the observation in real-time. The 'IUE observatory', thus comprises the ground control centre, where the astronomer views his television monitors, and the optical and electronic instrumentation located in synchronous orbit.

A synchronous orbit has other advantages. Because the Earth subtends an angle of only 17° the unconstrained area of sky is much greater than for low orbits and, as its movement along the ecliptic is at the diurnal rate, the occurrence of occultations is infrequent. Consequently, over large portions of the sky, neither

Table 1 Spacecraft characteristics

Characteristic	Description
Spacecraft mass	312 kg
Scientific instrument mass	122 kg
Apogee motor mass	237 kg
Launch vehicle adaptor mass	29 kg
Total launch mass	700 kg
Planned mission life	3-5 yr
Power required	210 W average
(Spacecraft and scientific instrument)	
Array capability (beginning of life)	424 W to 238 W depending on attitude
Launch vehicle	Delta 2914
Stabilisation and control	
Transfer orbit	Spin stabilised
Mission orbit	3-axis stabilised with better than 1 arc s control
Orbit	Elliptical geosynchronous
Inclination	28.6°
Apogee	45,887 km
Perigee	25,669 km
Eccentricity	0.24

long exposures nor the observation of variable phenomena need be periodically interrupted. A penalty is that in synchronous orbit observations are normally made in full sunlight so the telescope must be baffled to ensure adequate rejection of stray sunlight and earthlight. Fortunately, sunlight scattered from the Earth's atmosphere contains very little UV energy and so the limiting magnitude of the spectrograph is not appreciably affected by earthlight. The offset guiding system can be influenced, however, and relatively bright guide stars are needed near the Earth's limb.

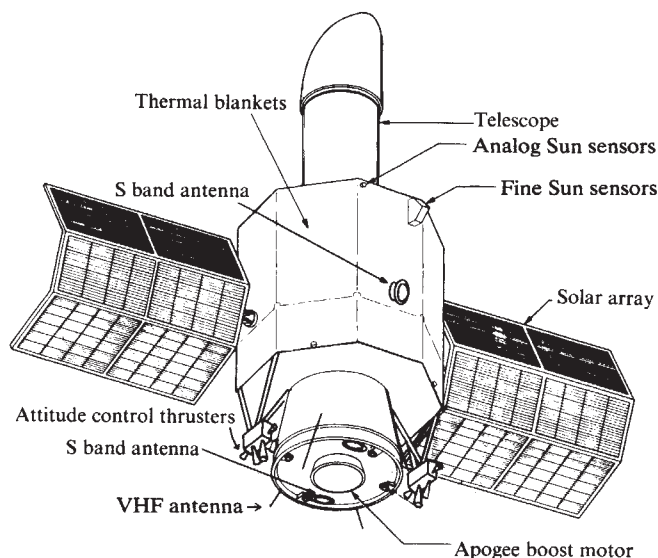


Fig. 1 IUE spacecraft.

General description of spacecraft

The general arrangement of the IUE spacecraft is shown in Fig. 1. The main body is octagonal in shape, the telescope protrudes from one end and the solar array extends from two opposite sides; in orbit the spacecraft is rotated around the telescope axis to keep the arrays nearly normal to the Sun. The apogee boost motor used to insert the spacecraft into synchronous orbit and the hydrazine auxiliary propulsion system used during the transfer orbit and for manoeuvres throughout the mission are both located in the lower cone assembly.

The telescope and spectrograph are rigidly connected together to form the scientific instrument to which the inertial reference assembly is directly attached to minimise relative displacement. An off-set star tracker within the scientific instrument provides the inertial reference assembly with a reference signal to enable drifts to be corrected. With this arrangement the pointing accuracy of the telescope is a fraction of an arc s.

Other spacecraft systems are described in later sections of this article and the spacecraft characteristics summarised in Table 1.

The scientific instrument

The scientific instrumentation consists of the telescope, the fine error sensors, the echelle spectrographs and the cameras. Each of these units will be discussed in turn.

Telescope and fine error sensor (FES): the telescope has an 0.45 m $f/15$ Cassegrain configuration with a Ritchey-Chrétien figure, giving uniform image shape across the 16 arc min field of view. The characteristics of the optics are summarised in Table 2. Telescope materials were chosen for their light weight, moderate cost, and thermal and optical properties. Beryllium was selected for the primary mirror, and fused silica for the secondary mirror. The telescope tube is aluminium, wrapped in aluminised Mylar thermal insulation; this thermal design produces a nearly constant focus over a wide range of solar aspect angles, and after initial focussing in orbit by movement of the

secondary mirror, no further refocussing should be required. Small corrections to the focus can be made using heaters behind the primary and secondary mirrors.

The forward sunshade is cut at a 43° angle with respect to the optical axis so that the telescope may be pointed anywhere in the sky more than 43° from the Sun. (The 43° angle was selected to permit observations of Venus when it is at greatest elongation.) The forward rim of the sunshade has an additional diffraction edge so that sunlight must be triply diffracted before it can enter the telescope tube. The small amount of diffracted sunlight that enters the telescope tube is further reduced by light baffles inside the tube. The most troublesome source of stray light is the Earth, which may appear at any angle with respect to the telescope axis. The baffles inside the sunshade are angled so that the light is preferentially scattered back out of the sunshade and all off-axis light must suffer at least two scatterings before it can reach a telescope mirror. With this design, stray light from the Earth or the Moon will be of negligible intensity as long as the primary mirror is not directly illuminated, but this can occur if the Earth's limb or the Moon is less than 25° from the optical axis.

Behind the primary mirror there is a shutter which closes automatically if sunlight falls inside the telescope tube. Behind the shutter, the spectrograph entrance apertures are mounted in the focal plane. These apertures are holes drilled in a mirror mounted at 45° to the telescope optical axis and which reflects the telescope field surrounding the apertures into duplicated FES, each of which serves the dual purpose of offset star tracker and field camera. Either is able to scan, in 8 arc s steps, the 16 arc min telescope field, an image of which can be transmitted to the ground for identification by the observer. The FES then measures the coordinates of the target star with respect to the spectrograph entrance apertures so that the spacecraft may be manoeuvred to centre the object in the selected entrance aperture. The FES is then offset to any other star in the field to generate error signals to hold the telescope on target. The limiting magnitude of the FES depends on its mode of operation, but it can track stars as faint as 14th visual mag. Thus any object, regardless of its brightness or nature, may be observed provided its position is known with respect to a nearby guide star.

Echelle spectrographs: behind the 45° aperture plate there are separate short and long wavelength echelle spectrographs, covering the spectral ranges 1,150–1,950 Å and 1,900–3,200 Å respectively. Each spectrograph has a pair of entrance apertures consisting of a hole 3 arc s in diameter and an oval slot 10×20 arc s. Both the larger slots can be closed simultaneously by a shutter, but the smaller holes always remain open. Any particular aperture can be selected by pointing the telescope so that the target image lies in that aperture. The arrangement of the optical components for the short wavelength spectrograph is shown in Fig. 2. Light from the target passes through the aperture plate to the collimator mirror, which is an off-axis paraboloid. The collimated beam is incident on the echelle grating and then diffracted to the spherical grating which acts

Fig. 2 Short wavelength spectrograph.

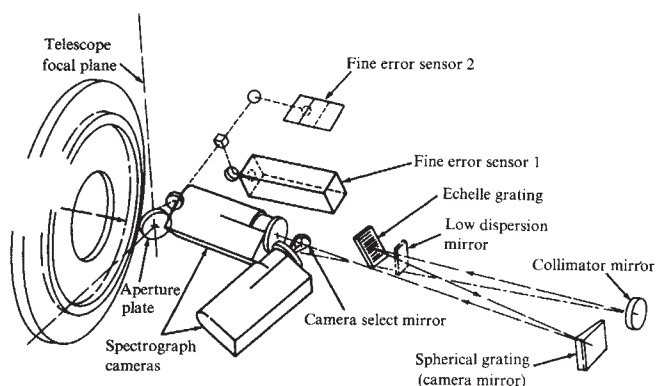


Table 2 IUE telescope parameters

Configuration	Cassegrain
Figure	Ritchey-Chrétien
Clear aperture	0.45 m
Central obscuration (baffled)	0.22 m
Primary focal length	1.25 m
Effective focal length	6.75 m
Focal ratio	15
Mirror separation	1.027 m
Back focal distance	0.175 m
Plate scale	30.6 arc s mm ⁻¹
Image quality	3 arc s
Field of view	16 arc min

both as an order separator and as a focusing element to image the entrance aperture on to the camera. By command, a camera-select mirror may be made to interrupt the beam and divert the image to a redundant camera. Similarly, a plane mirror (labelled the low dispersion mirror in Fig. 2) may be inserted in front of the echelle grating, resulting in a low dispersion spectrum produced by the spherical grating alone. The optical design of the long wavelength spectrograph is the same as the short wavelength spectrograph, except that two plane mirrors inclined at 45° are located behind the long wavelength entrance apertures to direct the incoming light beam to the long wavelength spectrograph and so separate its elements from those of the short wavelength instrument. In the long wavelength spectrograph the spherical grating could produce a second-order spectrum (from ~1,000 to 1,600 Å, which would overlay the 1,900–3,000 Å first-order spectral range), but this is suppressed by a silicon oxide coating on the 45° and collimator mirrors. The optical parameters of the two spectrographs are summarised in Table 3.

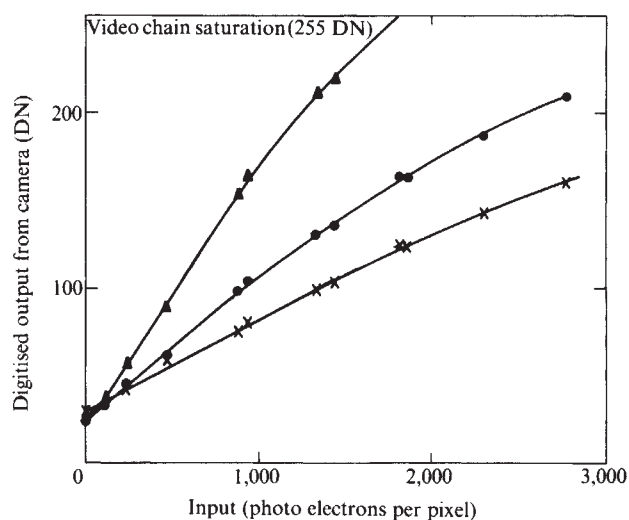
The spectrograph section of the scientific instrument also contains a number of light sources. The first of these is a platinum–neon hollow cathode lamp⁸ which is used as a wavelength standard. The lamp is mounted behind the primary mirror and illuminates the spectrograph entrance apertures by a mirror on the back of the Sun shutter so that a platinum spectrum can be recorded to calibrate wavelength as a function of position in the spectrograph image. Tungsten filament and mercury lamps are mounted in front of the cameras in each spectrograph to flood the camera faceplate with light. The former are used in the camera preparation procedure, described below; the latter⁹ can illuminate the cameras with predominantly 2,537 Å radiation, and are used to produce sets of flat-field images at different exposure times in order to derive intensity transfer functions for each picture element (pixel) of the camera format.

The spectrographs are fixed focus and, to minimise thermal effects, their temperature is closely controlled and the optics are mounted using tubes of graphite fibre-reinforced epoxy selected for its low coefficient of thermal expansion.

Spectrograph cameras: in each spectrograph a television camera integrates the two-dimensional echelle spectrum during an exposure and converts the image into video signals suitable for transmission to the ground. There are prime and redundant

cameras in both long and short wavelength spectrographs, all of nominally identical construction. Each camera head contains the electro-optical detector components and some of the electronics, such as the video head-amplifier, which must be close to the television camera tube. The remaining electronics are mounted in the spacecraft.

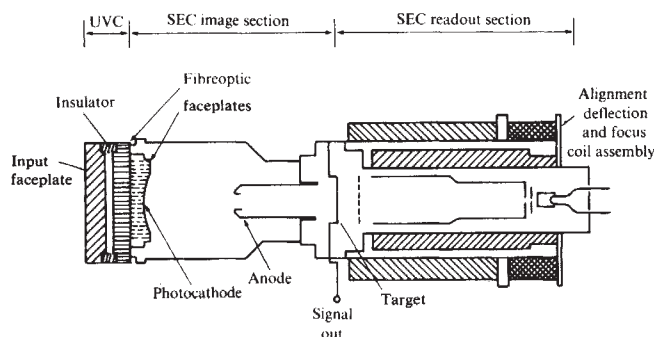
The basic detector is a UV-to-visible image converter (UVC) coupled to a secondary electron conduction (SEC) television camera which is sensitive to visible light only (Fig. 3). The entrance window of the UVC is magnesium fluoride, transparent to wavelengths above 1,150 Å, and on its inside face is a semi-transparent caesium–tellurium photocathode¹⁰ with a quantum efficiency of 10–15% in the UV but less than 0.01% in the visible. Within the UVC, a high electric field (5 kV across a 1.3 mm gap) accelerates photoelectrons onto a phosphor screen where each produces about 60 photons whose energy peaks in the blue. The screen efficiency is improved by a surface layer of aluminium which reflects backward-emitted photons through

**Fig. 4** Typical camera intensity transfer function for image centre (●) and positions of maximum (▲) and minimum sensitivity (×).

the output window; this layer also minimises optical feedback. A matt black coating on the input side of the aluminium reduces halation by suppressing reflection of UV photons which penetrate the photocathode.

The blue image from the UVC is transferred by fibre-optic coupling to the bi-alkali photocathode of the SEC tube. Incorporated in the fibre-optic coupling is a square array of fiducial marks to provide a geometrical reference for the spectral image. The photoelectrons, generated at the photocathode, are accelerated through the electrostatically-focused image section onto the SEC target which consists of a low density porous layer of potassium chloride about 10 μm thick, supported by an aluminium oxide membrane about 50 nm thick, and backed by a thin aluminium signal plate. Some fraction of the energy of each photoelectron is expended in producing several secondary electrons in the potassium chloride which are swept towards the signal plate by a 12V bias across the target leaving an amplified positive charge-image stored on the target, (the process of secondary electron conduction from which the SEC tube¹¹ is named). The gain depends on the accelerating voltage and is about 50 at the highest voltage setting of around 6.1 kV, because of the highly insulating nature of the target an image can be accumulated over many hours and there is no detectable reciprocity failure.

The image is read on completion of exposure. The readout section of the camera tube is similar to that of a standard magnetically-focused vidicon although its operation is somewhat different in that the beam is deflected digitally to scan a raster of 768 × 768 pixels in 37-μm steps across the image. At each pixel, the read beam is pulsed on and discharges the target,

Fig. 3 Camera showing UV to visible image converter and SEC tube.

giving rise to an output video pulse corresponding to the positive charge on the target at that pixel. The video signal is digitised into one of 256 discrete levels (DN units) by an 8 bit A/D converter, and these data are telemetered in real-time.

Table 3 IUE spectrograph optics

Optical element	Short wavelength spectrograph	Long wavelength spectrograph
Wavelength range	1,150 to 1,950 Å	1,900 to 3,200 Å
Offset mirrors	None	Two at 45°
Collimator radius	1.89 m	1.89 m
Low dispersion mirror	Plane	Plane
Echelle grating		
frequency	101.9 mm ⁻¹	63.2 mm ⁻¹
blaze angle	45.5°	48.1°
off-normal angle	10.2°	10.2°
Spherical grating		
frequency	313.0 mm ⁻¹	200.0 mm ⁻¹
radius	1.37 m	1.37 m
Camera select mirror	Plane	Plane
High dispersion resolving power*	1.2 × 10 ⁴	1.3 × 10 ⁴
Low dispersion resolution*	6 Å	8 Å

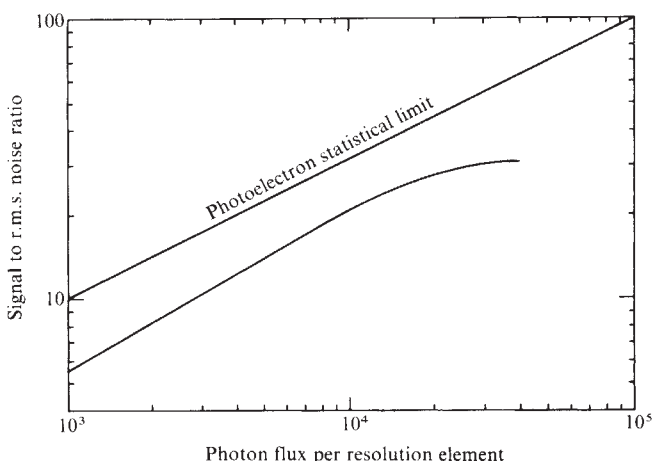
* Based on a predicted point spread function for the entire system (guidance plus optics plus detector) of 120 μm full width at half maximum. Actual resolutions achieved in orbit are discussed in the following paper¹.

The cameras are operated by a series of commands issued by the ground station operator through keyboard procedures. Each operational mode has a number of parameters which may be selected by the guest astronomer. The normal image sequence is: PREPARE-EXPOSE-READ. A PREPARE operation is carried out before each EXPOSE to erase all trace of previous images from the target and to provide a reproducible low-noise baseline on which the new image will be superimposed. It consists of preprogrammed sequences of over-exposures to the tungsten floodlamps followed by READ scans. In the EXPOSE mode the UVC and SEC image section high voltages are on (the readout section of the SEC tube is off) and an image may be integrated on the SEC target. Exposure times may be controlled from the ground or by the on-board computer. In the READ mode the image section and the UVC are off and the stored charge image is read from the SEC target as described above.

Camera performance

The response of a camera to UV photons is nonlinear at high signal levels, owing to saturation effects in the storage of charge in the SEC target. The relation between the number of input photoelectrons and the output from the video chain is known as the intensity transfer function (ITF) (Fig. 4). As can be seen, the ITF varies across the image and individual ITF corrections are applied to each of about half a million pixels. The photometric linearisation of the data is an important element of data reduction¹².

Fig. 5 Typical signal-to-r.m.s. noise ratio per resolution element (3 × 3 pixels) as a function of the input to the camera at 2,537 Å.



The ultimate accuracy with which a spectral feature can be measured is set by the quantum statistics of the signal photons incident on the detector but the quantum efficiency of the photocathode reduces this accuracy and it is further degraded by extraneous noise sources which depend on such parameters as SEC target gain, signal level, method of preparation of the target, wavelength and the camera noise level. A typical signal-to-noise characteristic achieved in orbit after photometric correction is shown in Fig. 5.

The spatial resolution of the camera is determined by the combined effects of the UVC and SEC tubes. A typical modulation transfer function (MTF) is shown in Fig. 6; the very rapid drop in MTF at low spatial frequencies is a result of halation.

The dark currents in the UVC and SEC are mainly thermionic in origin and result in a typical background signal of 0.5–1.5 DN per hour of exposure. Images resulting from long exposures usually show a number of small bright spots which vary in time and position; these are probably the result of radioactive disintegrations in the phosphor. Bright spots and streaks can also be caused by interaction of cosmic rays with the input window of the UVC. Aluminium shields not less than 4 mm in thickness surround each camera to provide substantial protection against radiation from the Earth's trapped electron belts which would otherwise produce an unacceptably high background.

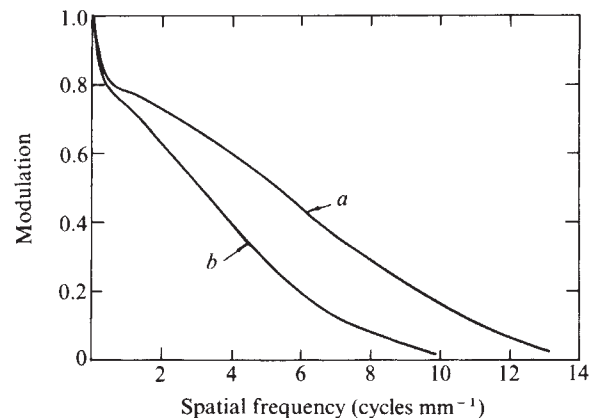


Fig. 6 Typical camera modulation transfer function for *a*, centre of image and *b*, at half radius (position of worst response).

Geometric distortions produced in the SEC tube amount typically to about 2% of the radial distance and consist of 'pin cushion' distortion produced by the electrostatically-focused image stage, 'S' distortion produced in the magnetically-focused read-out stage, small scale changes due to approximations used in the generation of the dynamic refocus waveform and small intensity-dependent image shifts in the read-out section of the SEC caused by 'beam-pulling' effects.

Spacecraft systems

The major sub-systems necessary to support operation of the scientific instrument are described below. Duplication is used extensively to ensure long term reliability.

Communications: the communications system consists of VHF transponders, S-band transmitters, r.f. power amplifiers and antennas. The characteristics of the VHF and S-band systems are summarised in Table 4. The S-band system is used only for transmission of telemetry data. The two transmitters can be connected to any of the four power amplifier-antenna combinations, but only one transmitter and one power amplifier may be selected at any one time and this will depend on which antenna has the most favourable view of the Earth. The VHF system consists of duplicate transponders and a four-element turnstile antenna and is used for the reception of ground-generated commands, the turn-around transmission of range and range-rate signals for tracking the spacecraft, and also to provide a backup telemetry down-link.

Table 4 Communications systems

	VHF	S-band
Transmitter frequency	136.860 MHz	2249.80 MHz
Power output	6 W	6 W
Modulation	PCM/FSK/AM	PM
Telemetry rate	800 b s ⁻¹	1.25 kb s ⁻¹ to 40 kb s ⁻¹ with fixed and reprogrammable formats
Antenna polarisation	Turnstile	Circular
Antenna pattern	Omnidirectional	60° conical
Receiver frequency	148.980 MHz	—
Receiver sensitivity	−106 db m	—

Commands initiated by the on-board computer or received directly from the ground are all processed by the two command decoders. Each can process 128 discrete commands and 48, 37-bit serial word commands.

Data handling: the data handling system is composed of the data multiplexer and the on-board computer. The data multiplexer serves as the spacecraft telemetry encoder and as the input data interface between the on-board computer and the rest of the spacecraft. 8-bit words are transferred to a serial data stream which is alternately made available to the ground and to the on-board computer using time sharing techniques. The telemetry bit rate is selectable by ground command from 1.25 to 40 kb s⁻¹. The on-board computer performs all attitude control computations and issues all reaction wheel torquing commands. It performs self-checks, monitors spacecraft performance safety functions, controls camera exposure times, and stores commands.

Stabilisation and control: the stabilisation and control system is used in orbital manoeuvres which include station keeping, pointing, and manoeuvring of the spacecraft. It has three different reference systems: (1) six gas-bearing, pulse-rebalanced, inertial-grade gyroscopes providing 0.01 arc s integrated rate resolution over 600 arc s per s; (2) star trackers (the FESs) in the scientific instrument which use the telescope optics to provide an angular resolution of 0.27 arc s throughout a 16 arc min field-of-view; (3) a two-axis digital Sun sensor which provides angular resolution to 15 arc s over a field-of-view of 64° × 124°. When a bright star is within 9 arc min of the target source, the guidance system uses the FES for position information and gyro system for rate damping. When a guide star is not available, precision-hold depends solely on a well-trimmed gyro reference with low frequency updates from a faint source.

Sequential three-axis manoeuvres from one star to the next use an integrated gyro reference resulting in accumulated errors of <4 arc min. The control system utilises a set of momentum-exchange reaction wheels for attitude control and relies on the hydrazine thrusters for momentum dumping.

Power: power is provided by two solar arrays and a highly efficient distribution and regulating system. During eclipse, and other periods when demand exceeds solar array output, power is provided through a boost regulator from two 6 amp h nickel-cadmium batteries.

Thermal control: the active thermal control system utilises louvres, insulation, heat pipes and heaters. Multi-nodal analysis of a thermal model and solar simulation tests were used to prove the design. The spacecraft may be divided into four sections, each with unique thermal requirements.

(1) The hydrazine bay which contains the hydrazine propulsion system and apogee boost motor and uses a combination of thermal blankets and active heaters.

(2) The main spacecraft compartment which is maintained between 0 °C and +40 °C with the batteries below 30 °C and with the gyros in the inertial reference assembly internally controlled at 57° ± 1 °C. Heat pipes attached to the main equipment platform, together with the louvres, keep gradients to <5 °C.

(3) The telescope, where heaters maintain the primary mirror near 0 °C and the focus mechanism between 0 and +30 °C. The primary mirror is conductively isolated from the telescope structure to reduce axial and circumferential gradients.

(4) The spectrograph, where the cameras are maintained between 0 and +20 °C. Temperature changes and gradients are minimised to prevent thermally induced misalignment. Satisfactory temperature control is obtained by enclosing the spectrograph within an aluminium shield to limit its heat exchange with the main spacecraft compartment and by temperature control of the primary mirror through which most of spectrograph's heat is dissipated.

Ground observatory control systems

Unlike previous unmanned astronomy spacecraft, IUE is operated in real-time by guest observers who generally lack detailed knowledge of the complex spacecraft and ground systems. Ground operating procedures have, therefore, been designed to allow the observer's research programme to be accomplished by selection from a library of modular pre-programmed operating sequences.

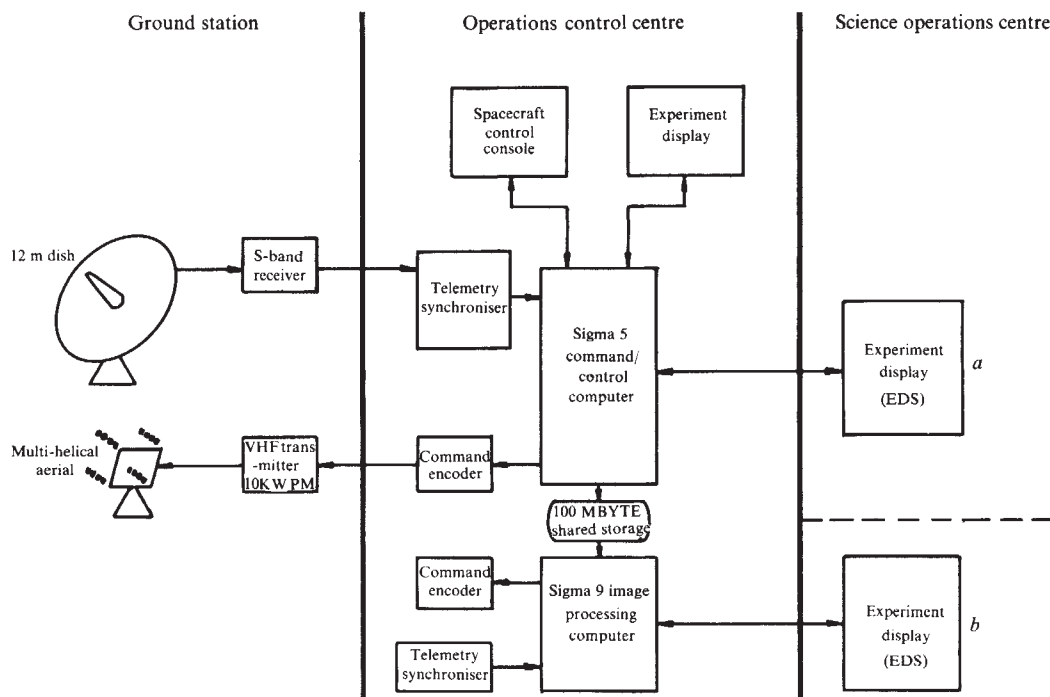


Fig. 7 US ground observatory. *a*, Observing centre; *b*, image processing centre.

IUE ground control is based on a large real-time computer software system to process telemetry and commands. Relatively complex spacecraft operations are accomplished by calling a series of operating procedures, each designed to accomplish a particular function, such as reading an image from a camera. Procedure execution is controlled by trained telescope operators. A computerised image processing system is then run offline to correct the astronomical images and produce a spectrum in absolute units as a function of wavelength¹³. The processing sequence consists of geometric and photometric correction, wavelength identification, data extraction and system efficiency correction and uses calibration tables derived from periodic analysis of calibration images.

The US ground observatory at GSFC is divided between three separate locations: the ground station, operations control centre, and science operations centre, see Fig. 7. The ground station maintains continuous radio communication with the spacecraft, relays telemetry data to the operations control centre and transmits command messages from the control centre to the spacecraft. The operations control centre contains a Xerox Data Systems (XDS) Sigma 5 computer used for real-time spacecraft operations and image acquisition, an XDS Sigma 9 computer used for image processing and data reduction, and a common disk storage unit. The Sigma 5 computer is operated from a spacecraft control console, consisting of a number of interactive

displays with keyboards used by engineers who continuously monitor the spacecraft performance. The science operations centre houses two experiment display systems (EDS) which are specially designed pseudo-colour image display and analysis consoles, containing their own mini-computers and disk storage. One EDS is used by the observer and telescope operator to conduct real-time observations and the other is used in image processing.

The ESA ground station and control centre is functionally identical with that at GSFC but incorporates a single Sigma 9 computer, with EDS, telemetry, command, and control consoles for operating the spacecraft. From this control centre the ESA- and SRC-sponsored observers use the IUE spacecraft for 8 h each day to conduct their research programmes.

1. Boggess, A. *et al.* *Nature* **275**, 361–369 (1978).
2. Heap, S. R. *et al.* *Nature* **275**, 369–372 (1978).
3. Linsky, J. L. *et al.* *Nature* **275**, 373–378 (1978).
4. Grewing, M. *et al.* *Nature* **275**, 378–384 (1978).
5. Dupree, A. K. *et al.* *Nature* **275**, 384–387 (1978).
6. Boksenberg, A. *et al.* *Nature* **275**, 388–398 (1978).
7. Lane, A. L. *et al.* *Nature* **275**, 398–399 (1978).
8. Mount, G. H., Yamasaki, G., Fowler, W. & Fastie, W. G. *Appl. Opt.* **16**, 591 (1977).
9. Boerio, A. H., Beyer, R. R. & Geotze, G. W. *Adv. Electr. Electron Phys.* **22A**, 229 (1966).
10. Coleman, C. I. *Appl. Opt.* **17**, 1789–1796 (1978).
11. Childs, C. B. *Appl. Opt.* **1**, 711–716 (1962).
12. Allen, K. G. R., Anderson, B. E., Boksenberg, A. & Oliver, M. B. *Adv. Electron. Electron Phys.* **40A**, 223–237 (1976).
13. Golton, E. & Dunford, E. *J. Br. interplanet. Soc.* (in the press).

In-flight performance of the IUE

A. Boggess, R. C. Bohlin, D. C. Evans, H. R. Freeman, T. R. Gull, S. R. Heap, D. A. Klinglesmith, G. R. Longanecker, W. Sparks & D. K. West

Goddard Space Flight Center, Greenbelt, Maryland 20771

A. V. Holm, P. M. Perry, F. H. Schiffer III, B. E. Turnrose & C. C. Wu

Computer Sciences Corporation, Silver Spring, Maryland 20900

A. L. Lane

Jet Propulsion Laboratory, Pasadena, California 91103

J. L. Linsky

Joint Institute for Laboratory Astrophysics, Quantum Physics Division, NBS and University of Colorado, Boulder, Colorado 80309

B. D. Savage

University of Wisconsin, Madison, Wisconsin 54220

P. Benvenuti, A. Cassatella, J. Clavel, A. Heck, F. Macchetto, M. V. Penston & P. L. Selvelli

ESA Satellite Tracking Station, Villafranca Del Castillo, Madrid, Spain

E. Dunford, P. Gondhalekar, M. B. Oliver, M. C. W. Sandford & D. Stickland

SRC Appleton Laboratory, Slough, Berks, UK

A. Boksenberg, C. I. Coleman, M. A. J. Snijders & R. Wilson

University College London, Gower Street, London WC1, UK

After the IUE's successful launch into an eccentric geosynchronous orbit, and following initial spacecraft checkout, observations were made of a set of high priority targets as an insurance against premature failure of the system. These were followed by the systematic performance evaluation of the various spacecraft and scientific instrument sub-systems, some of which were optimised in orbit. The optical performance of the telescope and spectrograph and the photometric performance of the SEC cameras used as spectrum detectors all seem satisfactory, but the ground data processing system although operational needs further development.

THE IUE was launched from the John F. Kennedy Space Center on 26 January 1978 into an eccentric geosynchronous orbit. The system is described in the previous article. The mission objectives required that the spacecraft be placed in a 24-hour orbit visible continuously from the NASA ground station at Goddard Space Flight Center and for at least 10 h per day from the ESA Villafranca Satellite Tracking Station near Madrid. This coverage allows the Goddard Control Center to control operations for 16 h each day and to monitor health and safety of the spacecraft while the Villafranca Tracking Station operates the satellite for 8 h each day. The orbital elements are summarised in Table 1. The satellite tends to drift slowly in longitude because of slight inaccuracies in the orbital period. However, the ascending node is kept near the longitude indicated in Table 1 by

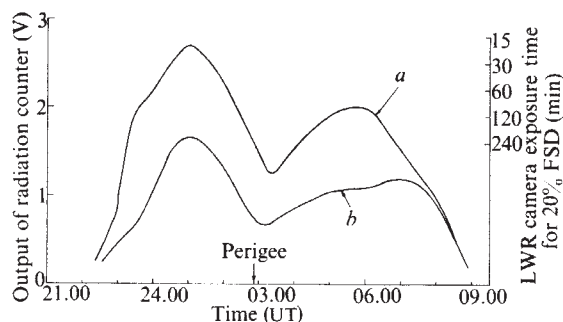


Fig. 1 Typical variation of electron flux (≥ 1 MeV) measured by the on-board counter. The counter output voltage, V , is related to the electron flux, F (electrons $\text{cm}^{-2} \text{s}^{-1}$), by $V = \log_{10} F - k$ where $k \approx 3.4$. The right hand scale gives the exposure time in minutes that would give 20% of full scale on the long wavelength redundant camera. This calibration was derived at several counter levels between 1.5 and 2.5 V. The short wavelength prime camera is 1.4 times more sensitive to radiation. *a*, 31 January–1 February; *b*, 30–31 January.

occasionally firing hydrazine thrusters. These small orbital adjustments will be required two to three times per year.

After the spacecraft was injected into the synchronous orbit, three-axis attitude control was established and the integrity of the basic spacecraft and scientific instrument systems was verified. At this point a selected set of high priority targets was observed so that some scientific data spanning various research areas would exist in the event of an early failure. An IUE science commissioning team, including project personnel and a few guest observers, had been formed to direct the early scientific checkout, observing, and performance evaluation activities. Each team member was supported by one or more associates so that the team had a large staff of scientists and engineers to assist in the commissioning tasks. The tasks covered three basic phases. (1) Immediately after turn-on, the selected high priority targets were observed to develop observing techniques and to collect some data in each of the broad research categories applicable to IUE. The observations are described in the following papers²⁻⁷ and fall into the categories of hot stars², cool stars³, interstellar medium⁴, X-ray objects⁵, extra-galactic objects⁶, and solar system objects⁷. (2) The scientific instrument and supporting spacecraft equipment were optimised. During this phase the telescope was focused, prime and back-up cameras were selected for each spectrograph, the settings of each prime camera were adjusted to achieve best performance, and the slewing and guiding capabilities of the spacecraft were improved. (3) The photometric properties of the scientific instrument were calibrated against standard stars, and other performance characteristics, such as image quality and spectral resolution, were studied.

At the end of these three phases routine guest observing began on 3 April 1978. The optimisation, evaluation, and calibration tasks were far too large to be fully completed within the available time. However, those items necessary for starting the routine observing phase were finished, and the basic procedures established whereby the scientific commissioning will be completed and maintained by the observatory personnel.

Spacecraft sub-systems

The spacecraft sub-systems whose proper operation are essential to the mission are described in the previous article¹. They include power, thermal, communications, data handling, and attitude control, and all are operating successfully. The power sub-system has performed according to its design since the launch. The solar arrays totally support normal operations over the acceptable range of β -angles (the angle between the telescope axis and the direction to the anti-Sun) from 0 to 135°. As the arrays degrade due to radiation damage the β -angle range for positive power will decrease, and the power system will begin to depend on stored energy from the batteries to maintain operation over all β angles. The batteries and power system

performed normally during the first eclipse period, during which the spacecraft passed through the Earth's shadow every day for about three weeks. Eclipse seasons for IUE, as for the Moon, occur every six months.

The overall thermal performance of the IUE spacecraft has been satisfactory. The only departures from pre-flight predictions are that the power sub-system electronics are running warmer than expected, and there is a larger than anticipated gradient between the two battery packs. There was some initial concern about how the batteries would perform under load with this gradient; however, no problems arose during the first eclipse period. The temperatures throughout the scientific instrument are close to the predictions and the scientific instrument heaters are routinely cycled to keep its thermal conditions within the desired range.

The spacecraft communication and data-handling sub-system has been used in all its various formats and bit rate configurations, and it operates flawlessly. Most operations use a telemetry rate of 20 kbits s^{-1} . The command sub-system is performing properly with hundreds of commands issued to the spacecraft each day to support routine real-time operations. Occasional command-verify anomalies have been observed, but in most cases these have been traced to ground transmitter adjustment errors. Occasional interference is encountered from other NASA satellites during ranging measurements. The primary on-board computer has been in operation since spacecraft turn-on. During the first few weeks several instruction failures were observed which occasionally caused the computer to halt. However, software modifications were installed to detect such conditions and recover automatically without impact on scientific operations. While the primary function of the computer is attitude control, its other functions include timing camera exposures and monitoring certain spacecraft parameters. As with the other spacecraft systems, a redundant on-board computer is available in case the prime system should fail.

The attitude control sub-system operates satisfactorily in all modes of operation. It controlled nutation and precession while the spacecraft was spinning in the transfer orbit and, after injection into the synchronous orbit, it executed the de-spin and Sun acquisition manoeuvres. Three-axis control was established using the on-board computer, the inertial reference assembly, and the reaction wheels. The initial attitude reference was obtained by measuring the solar position with fine Sun sensors in pitch and roll and the approximate Earth position with a panoramic attitude sensor in yaw. The spacecraft was then slewed to point the telescope toward the North Ecliptic Pole, where a star field was identified using the telescope and star tracker. Manoeuvres to new positions are performed, one axis at a time, by inserting in the on-board computer new gyro references relative to the known spacecraft orientation. Reaction wheels slew the spacecraft, and hydrazine reaction jets are fired once or twice daily to reduce wheel momentum that accumulates due to disturbance torques on the spacecraft. Typical pointing errors after large angle slews have been 1–2 arc min, which is well within the 16 arc-min field of the telescope. The star tracker in the scientific instrument can identify and guide on stars brighter than 14 mag. Using the star tracker for guidance, pointing accuracies ~ 0.2 arc s have been achieved. On the rare occasions when there is no suitable guide star in the field, the gyro drift rates may be nulled while the tracker is guiding on a nearby star and then the gyros may be

Table 1 Initial orbital elements

Semi-major axis	42,152 km
Perigee	32,050 km
Apogee	52,254 km
Eccentricity	0.240
Inclination	28.6°
Right ascension, ascending node	208.1°
Argument of perigee	257.1°
Period	23h55min33s

LWR camera exposure time (min) for 20% FSD

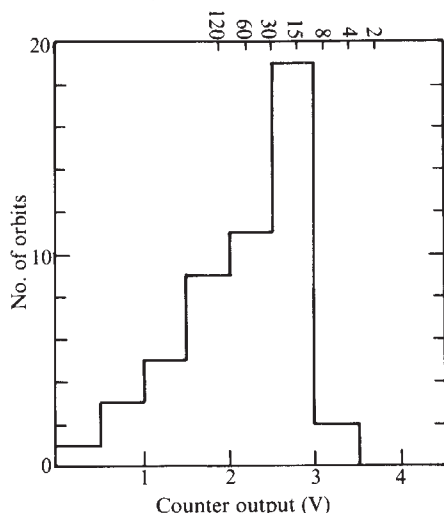


Fig. 2 Histogram showing the variation in peak flux during each of 50 orbits for which data are available for the period 29 January–10 April 1978.

used for inertial guidance with random drifts ~ 3 arc s per h. If moving targets are to be observed, the gyro rates can be adjusted to track the target's motion.

Optical performance

Measurements of telescope image quality before launch indicated that the ratio of transmission between the small spectrograph entrance aperture (3 arc s) and the large spectrograph entrance aperture (10×20 arc s) was 0.65 at 20°C , 0.60 at 5°C , and 0.55 at -10°C . Because of air turbulence and the angular diameter of the laboratory source, these measurements were considered to be lower limits of achievable performance.

Thermal conditions have been adjusted in orbit and are being controlled actively to maintain a nominal primary mirror temperature of -3°C . This should allow pointing to any object outside the 43° Sun avoidance cone without significant changes to telescope focus. After refocusing the telescope in orbit, spectrograms of the same star obtained using the large and small apertures in the short wavelength spectrograph have indicated small aperture transmissions of up to 70%. Spectra of stellar images in the large aperture are ~ 5 arc s full width at half maximum, perpendicular to dispersion.

Sources of scattered light in the telescope may be divided into three categories: the Sun; Earth and Moon; and bright stars. The IUE telescope baffles (see ref. 1) were designed to eliminate sunlight over the region of sky outside the 43° half-angle Sun avoidance cone. No evidence of scattered sunlight has been observed at angles of 50° or more from the Sun. At angles between 45° and 50° , stray sunlight is barely detectable in the star trackers. There are no restrictions on observing near the Earth or Moon, but scattered light may affect the operation of the star trackers. Observations of stars with no evidence of scattered light contamination have been made within 25° of the Moon and within 25° of the Earth's limb. The practical limit has not yet been fully determined, but when the Earth is within the 16 arc min field-of-view, the star tracker output is saturated in every operational mode. Observations of Jupiter and its satellites have shown that 6th mag objects can be observed and tracked within 90 arc s of a -1.7 mag object. This performance would scale to tracking a 14th mag object at 90 arc s from a 6.3 mag object. It is also important to know how severely the spectrum of an object can be contaminated by scattered light from a bright star located close to it. Tests on η UMa show attenuations of 10 stellar magnitudes when the star is 10 arc s from the small entrance aperture of the long wavelength spectrograph and 12 stellar magnitudes at 20 arc s.

The resolutions of the long wavelength and short wavelength spectrographs have been determined from spectra of the Pt–Ne

hollow cathode lamp that is used on-board for wavelength calibration. At high dispersion in the short wavelength spectrograph, the full width at half maximum (FWHM) of a spectral line from the Pt–Ne lamp is $0.085 \text{ \AA}$ at $1,150 \text{ \AA}$, and it increases linearly with wavelength to $0.19 \text{ \AA}$ at $2,100 \text{ \AA}$. In the long wavelength spectrograph, $\text{FWHM} = 0.20 \pm 0.01 \text{ \AA}$, constant with wavelength. The data for the short wavelength spectrograph correspond to a resolving power, $R = \lambda/\text{FWHM} \approx 1.2 \times 10^4$, as was expected. The data for the long wavelength spectrograph give resolving powers that vary from 1.0×10^4 to 1.5×10^4 over the wavelength range rather than the approximately constant value of 1.3×10^4 that was expected. The reason for this is not yet understood. The spectral resolution measured from Pt–Ne spectra in the low dispersion modes is $\leq 7 \text{ \AA}$ in the short wavelength spectrograph and $\leq 8 \text{ \AA}$ in the long wavelength spectrograph. These results are upper limits because all features of the Pt–Ne spectrum observed at low dispersion are, to some extent, blends of several lines.

The on-board Pt–Ne lamp⁸ serves two purposes. First, it defines the spectral format to guide the extraction of data from the spectrum and its local background; this is clearly a more intricate process in the high dispersion mode with its 2-dimensional format. Second, it assigns a wavelength to each data point of an extracted spectrum. The procedure involves two steps. First, the emission line spectrum of the lamp is obtained, and the positions of the spectral lines on that image are found using a cross-correlation technique. In the case of the high dispersion spectra ~ 200 spectral lines are located, while in the case of low dispersion spectra the number is ~ 15 – 20 . Second, a multivariable regression analysis is performed to relate the sample s and line l positions to their wavelength λ and order numbers m . The equations of condition are given below for high dispersion:

$$s = a_0 + a_1 m \lambda + a_2 (m \lambda)^2 + a_3 m + a_4 \lambda + a_5 m \lambda^2 + a_6 m^2 \lambda$$

$$l = b_0 + b_1 m \lambda + b_2 (m \lambda)^2 + b_3 m + b_4 \lambda + b_5 m \lambda^2 + b_6 m^2 \lambda$$

and for low dispersion

$$s = a_0 + a_1 \lambda \quad l = b_0 + b_1 \lambda$$

These dispersion constants are then used to define the spectral format of the target spectrum. The errors in this procedure are assessed in Table 2, which shows that the largest uncertainty in assigned wavelengths arises in transferring the dispersion constants derived from a Pt–Ne spectrum to the image of a stellar spectrum. These errors can be removed by identifying a feature in the astronomical source and applying an appropriate correction.

Table 2 Sources of error in wavelengths

Source	High dispersion	Low dispersion
Error in laboratory wavelength	$<0.004 \text{ \AA}$	$<1 \text{ \AA}$
Error in measured positions of spectral lines	$<0.07 \text{ \AA}$	$2 \text{ \AA}$
Inadequacy of dispersion formulae to represent spectral format	$<0.07 \text{ \AA}$	$<2 \text{ \AA}$
Error in transferring dispersion relations to target spectra	$0.2 \text{ \AA}$	$6 \text{ \AA}$

Camera performance

Before launch an extensive programme was undertaken to investigate the effect on camera performance of the settings of the camera-tube and deflection-coil operating parameters⁹. This programme resulted in a set of values for the commandable voltages that was used immediately after launch to observe the high priority targets. After the high priority observing period, time was allocated to check the commandable camera settings, using the performance before launch as a baseline. In addition, the post-launch check on the operating parameters was extended to include the implementation of a new target preparation sequence, called SPREP, which had been developed during additional laboratory testing of the flight spare cameras.

Table 3 Camera performance

Camera	Photometrically corrected 40% image				Random noise on pedestal			Preparation sequence
	Mean S/N* (Single pixel)	Min. S/N (Single pixel)	Mean S/N (3 × 3 pixels)	Photometric† stability	Mean σ ‡ (Single pixel)	Peak σ (Single pixel)	Mean σ (3 × 3 pixels)	
Long wavelength redundant	10.0	5.8	21.2	±7%	3.0 DN	7.9 DN	1.1 DN	TPREP: ground calibration
	13.4	8.0	29.0	±2.5	1.4	2.7	0.4	SPREP: after re-optimisation
	8.7	5.5	19.4	±5.5	2.9	6.0	1.0	SPREP: in-orbit ITF
	7.8	3.5	19.2	±7.5	2.7	5.7	0.9	FPREP: after re-optimisation
Short wavelength prime	12.2	4.0	30.2	±13.5	1.6	2.8	0.7	TPREP: ground calibration
	10.5	6.5	23.1	±2.5	1.4	2.9	0.5	SPREP: after re-optimisation
	10.9	6.6	25.0	±3.2	1.7	3.6	0.6	SPREP: in-orbit ITF
	11.1	6.0	25.0	±5	1.3	2.2	0.6	FPREP: after re-optimisation

All results refer to the low gain setting of the head amplifier with the SEC tube at maximum gain.

* The mean signal-to-noise (S/N) ratio is averaged over the whole target area for both individual pixels and for 3 × 3 pixel blocks. The minimum S/N ratio is the mean over a block of 24 × 24 pixels in the area of worst noise.

† The photometric stability represents the peak-to-peak spread in signal level after photometric correction has been performed.

‡ The mean standard deviation (σ) is averaged over the whole target area for both individual pixels and for 3 × 3 pixel blocks. The peak standard deviation is the mean over a block of 24 × 24 pixels in the area of worst noise performance.

SPREP has now replaced the original preparation sequence, TPREP, which was used during high priority observing. Further images were obtained to serve as a new baseline for measuring long term drifts in camera performance. Two other target preparation sequences, also initially developed in laboratory testing, have been implemented. These are FPREP, a fast preparation sequence designed for use when speed of operation is important and some degradation in image quality can be tolerated, and XPREP, which can be used to remove residual effects of over-exposed images.

The function of the preparation sequences is to remove all trace of previous images and to leave the target with a uniform low-level low-noise pedestal. Target preparation involves a saturation exposure of the camera, using tungsten floodlamps, and read out of the charge, followed in the case of TPREP and SPREP by a lower level flood exposure and subsequent read. The original preparation sequence, TPREP, has the disadvantage that beam pulling effects can lead to small scale instabilities in the action of the read beam, particularly when reading out the saturated exposure of the preparation sequence. These beam instabilities result in patches of increased noise and poor photometric stability. The new sequence, SPREP, gives significant improvements in photometric stability and considerably reduced noise on low-level signals. The improvements are

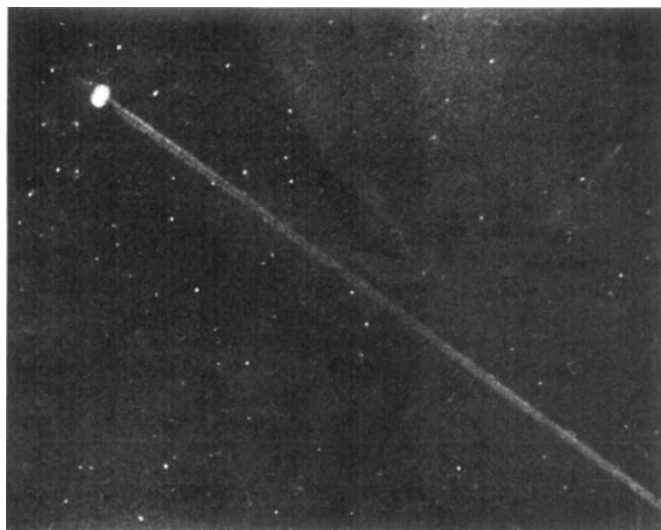
achieved principally by scanning the target with a defocused beam and with carefully chosen alignment settings, permitting a much more stable scanning action. Other improvements in noise performance are obtained by making the second floodlamp exposure at a lower gain setting, thus reducing the effect of photon noise, and by increasing the level of this exposure from 20% of saturation to 50%, at which level the random noise on the pedestal has been shown empirically to pass through a minimum value.

Table 3 lists the camera performance before and after reoptimisation. The performance using SPREP is significantly better than that for TPREP. The photometric stability is considerably improved, the signal-to-noise ratio in the worst areas has been improved, and the random noise on the pedestal has been reduced, particularly in the case of the long wavelength redundant camera. Also shown in Table 3 is the camera performance derived from the photometric calibration images taken shortly after the end of the reoptimisation programme. The performance of the short wavelength prime camera is comparable with that determined from the evaluation images taken as part of the reoptimisation work. However, in the case of the long wavelength redundant camera the performance seems to have degraded. The reason is not yet fully understood and more analysis remains to be completed. The apparent change may be a result of temperature transitions in the region of the detector or associated electronics which could cause a slight shift in the position of individual pixels on the scanning raster and hinder the removal of fixed pattern noise.

FPREP, the fast target preparation sequence, is achieved using a single floodlamp exposure and carrying out the read operations at a scanning rate of 320,000 pixels s⁻¹, which is 64 times the speed of the fastest scanning rate used in other modes and is considerably faster than the response time of the line deflection digital-to-analogue converter. The beam is not pulsed as it is in normal operation, but is held on continuously during the scan and the staircase waveform applied to the deflection coil is smoothed into a slope on which are superimposed spikes resulting from 'glitches' in the digital-to-analogue converter. These spikes give rise to a series of vertical bars, typically about 10% of saturation, superimposed on the image and degrading the quality of the data. It had been thought that these bars would not be stable with time, preventing their removal by means of photometric correction, and Table 4 shows that FPREP is significantly inferior to SPREP in the case of the long wavelength redundant camera, though still quite adequate for its intended use. The performance of FPREP for the short wavelength prime camera compares favourably with that using SPREP. It remains to be seen whether the bars are stable in the long term.

XPREP, which is used to remove the residual components of highly over-exposed images, is similar to FPREP, except that the floodlamp exposure level is four times larger, driving the

Fig. 3 Contrast-enhanced but otherwise unprocessed low-dispersion short wavelength spectrum of the faint ($B \sim 14$ mag) extragalactic object B2 1101+38 exposed for 4 h. The geocoronal Ly α emission fills the 10 × 20 arc s aperture and can be distinguished from the narrower object spectrum. There is no sign of background from other atmospheric emissions either in this or the long wavelength spectrum. The numerous bright spots are generated in the detector. The comet-like image near top right is a cosmic-ray induced event occurring in the detector optical components.



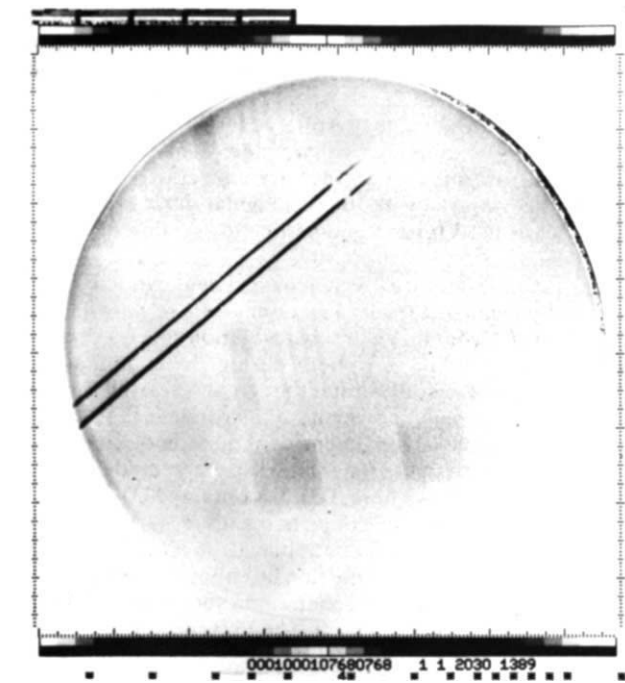


Fig. 4 Photographic representations of a raw camera image showing two low resolution short wavelength spectra of θ^2 A Orionis (O9.5 Vp, $V = 5.07$). The upper spectrum was exposed for 1 s through the large aperture and the lower spectrum was exposed for 4 s through the small aperture. The spectra extend from $\sim 2,000$ to $1,150$ Å. The strong interstellar Ly α absorption line at $1,216$ Å is clearly seen at the short wavelength end of each spectrum.

target well into saturation and completely obliterating any residual images. XPREP is normally followed by SPREP in order to obtain a reproducible low-noise pedestal. It has been used on both the long and short wavelength cameras following exposures containing spectral lines that were over-exposed by factors of up to 1,000. Initial results show that all traces of residual image are removed, while control exposures using SPREP only show residual signal levels of several per cent in line images that were over-exposed by factors of $\sim 1,000$.

The efficiency of IUE depends on the amount of overhead time associated with each exposure. In addition to the times for slewing and for target acquisition there are overheads associated with the operation of the cameras. As shown in Table 4 these fall into the categories of camera hardware overheads and ground control computer overheads. Considerable improvements were made to the latter during commissioning and further improvements are expected in the procedures by which the computer controls camera operations. Table 4 shows overheads for the two most commonly used telemetry rates, 20 and 40 kbits s^{-1} . It is often possible to hide the overhead associated with SPREP and READ by performing them during a long exposure in the other spectrograph; SPREP may also be carried out during a slew.

Background corrections

There are several contributors to the background which are apparent in camera images. The first of these is the null pedestal discussed in the previous section. This pedestal is a zero offset, typically amounting to 15% of saturation, that must be sub-

Table 4 Overhead time associated with camera operations

Operation	Execution of procedure in ground computer	Camera hardware time		Total overhead	
		at 20 kbit s^{-1}	at 40 kbit s^{-1}	at 20 kbit s^{-1}	at 40 kbit s^{-1}
SPREP	10 m	13 m	7 m	23 m	17 m
FPREP or XPREP	4 m	1 m	1 m	5 m	5 m
EXPOSE	4 m	0 m	0 m	4 m	4 m
READ	10 m	5.5 m	3 m	15 m	13 m

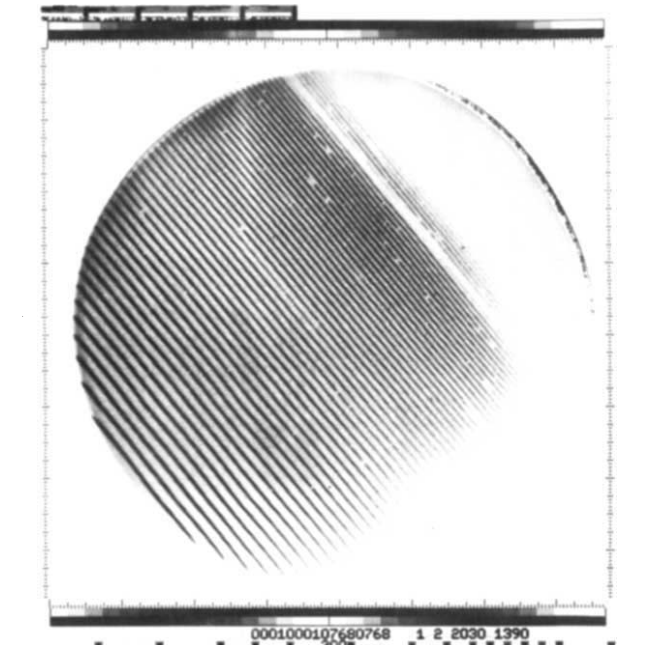
Table 5 Sample 1- σ noise in extracted spectra				
Camera	Object	Image	σ_{meas} (%)	σ_Q (%)
1	ζ Oph	LWP1031	4.2	1.9
2	η UMa	LWR1172	2.8	2.3
3	η UMa	SWP1164	<2.4	2.1
4	ζ Oph	SWR1031	3.6	1.8

tracted from each image. Although there is in principle a time-dependent internal tube noise due to thermal effects and to phenomena such as phosphorescence in the image converter, in practice the tube background appears to be negligible for exposures as long as 5.5 h, which is the longest integration time attempted so far.

During the passage of the satellite through the Earth's outer belt of trapped electrons there is often a significant background level in the data due to Cerenkov radiation in the input window. There is apparently no effect on the stored image in the SEC tube when the camera is in the STANDBY or READ modes. A silicon diode ionisation counter on board the spacecraft serves as a real time monitor of the flux of electrons ≥ 1 Mev, which are those expected to penetrate the radiation shielding and produce Cerenkov radiation. Figure 1 shows the variation of the output of the counter during parts of two typical orbits. The main maximum always occurs before perigee as the satellite passes just above the peak of the outer belt on its way south. The second maximum occurs as the satellite passes north at a greater altitude. The result of the in-orbit calibration of the response of the cameras to this radiation is also shown in Fig. 1. Using this calibration and the real time output from the radiation counter the observer can arrange to terminate his exposure at a suitable background level.

Figure 2 shows the statistics of the day-to-day variation in maximum radiation flux. From this figure it will be seen that for $\sim 40\%$ of the time exposures of over 2 h are possible during the passage through the belt. During the remaining 60% of the orbits, there will be restrictions on the maximum exposure time for periods between 2 and 8 hours duration. The radiation level in the belt is expected to increase as the peak of solar activity is approached.

Fig. 5 Photographic representation of a raw camera image showing a high resolution short wavelength spectrum of θ^2 A Orionis. The exposure was for 5 min through the small aperture. Echelle orders in the lower left portion of the figure provide wavelength coverage to $\sim 2,100$ Å while orders in the upper right provide wavelengths down to $\sim 1,150$ Å. The broad interstellar Ly α absorption line at $1,216$ Å nearly eliminates two complete echelle orders. Numerous narrow interstellar lines can be seen but they should not be confused with the rectangular pattern of fiducial marks.



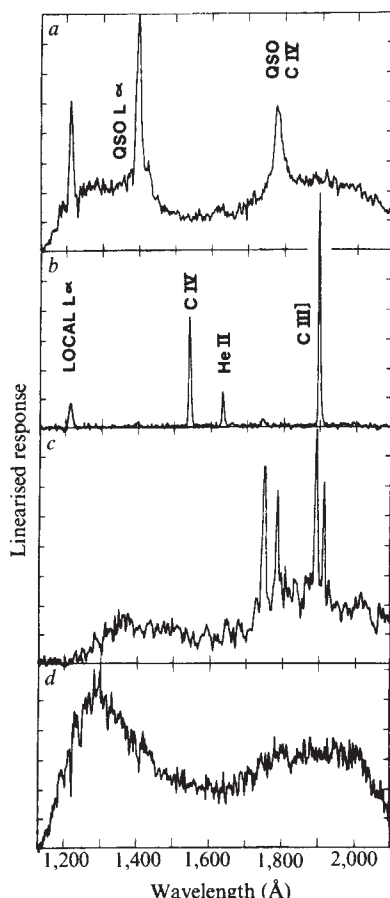


Fig. 6 A selection of low dispersion short wavelength spectra obtained by IUE during the commissioning period. Some of the strongest lines are identified. *a*, 3C273; *b*, NGC7027; *c*, η Car; *d*, BD +75°325.

The only detectable sky background source is geocoronal Ly α emission. This is evident in the example shown in Figure 3, which is a contrast-enhanced, low dispersion spectrum taken with the short wavelength redundant camera of the faint ($B \sim 14$ mag) extragalactic object B2 1101+38 exposed for 4 hours during the commissioning phase. Here the geocoronal Ly α emission, which fills the 10×20 arc s entrance aperture, can be distinguished easily from the object spectrum, which appears as a point source. There is no sign of background from atmospheric continuum emission or from other line emissions, such as the O I multiplet near λ 1304 or [O I] λ 1356, which are known to peak far below the satellite altitude. There are some bright spots generated in the detectors themselves; residual optical 'hot spots' may fluoresce in the phosphor of the UV converter following the flooding of the preparation procedure, and phosphor radioactive decay spots and internal ion events appear occasionally. Finally, cosmic-ray induced events, which exhibit a comet-like appearance, may occur in the detector optical components. These are distinct from the diffuse Cerenkov radiation background caused by trapped radiation. Cosmic rays usually produce small clusters of saturated pixels that must be ignored in analysis of the data. Clearly, it is important to use the basic pictorial data to distinguish real from spurious lines by verifying that any apparent feature is the same width as the actual spectral image.

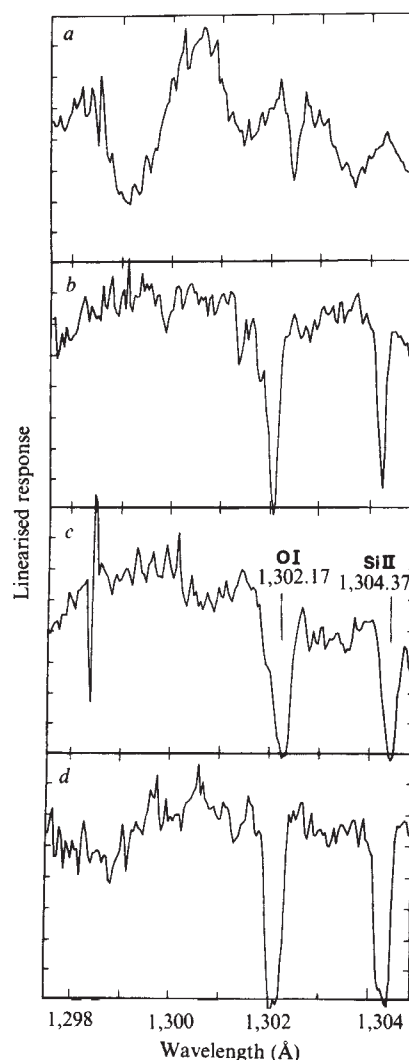
Photometric performance

Figures 4 and 5 show representative raw images obtained with the short wavelength prime camera at low and high dispersion. These are photographic representations of the original camera images and are a standard data reduction product. For both figures the object is θ^2 A Orionis (O9.5 Vp, $V = 5.07$), a bright early type star in the vicinity of the Orion nebula. For the low

resolution exposure in Fig. 4, spectra were obtained through both entrance apertures and the resultant spectra extend from $\sim 2,000$ to $1,150$ Å. At the short wavelength end the interstellar Ly α line at $1,216$ Å is clearly seen. For the high dispersion spectrum in Fig. 5 the small aperture was used. The interstellar Ly α line nearly eliminates two complete echelle orders and numerous narrow interstellar lines are also clearly visible but should not be confused with the rectangular pattern of fiducial marks which is used for the geometrical correction of the raw images.

Several sample spectra, which have been extracted from camera images similar to those just discussed, are shown in Figs 6 and 7. Figure 6 contains four low resolution spectra obtained with the short wavelength camera during the commissioning period. These spectra illustrate the great versatility of the instrument in observing a diversity of astronomical objects. A dynamic range of 18 mag in the incident flux can be covered by using appropriate exposure times and dispersion mode. In addition to the samples shown here, IUE has obtained UV spectra of cool stars, other extragalactic objects, standard stars, and Solar-System objects, including Mars, Jupiter, Uranus, and the Moon. Spectra ranging from a pure emission line object in the planetary nebula NGC7027 to a smooth continuum source in the hot star BD +75°325 are shown in Fig. 6. The presence of some noise due to quantum statistics and residual system noise is evident,

Fig. 7 A selection of high dispersion stellar spectra from order 106 of the short wavelength echelle spectrograph. The two strong absorption lines of O I and Si II are primarily interstellar. The glitch on the left in HD153919 is produced by one of the camera fiducial marks which are about 3 pixels wide. *a*, η UMa; *b*, ζ Oph; *c*, HD153919; *d*, HD93521.



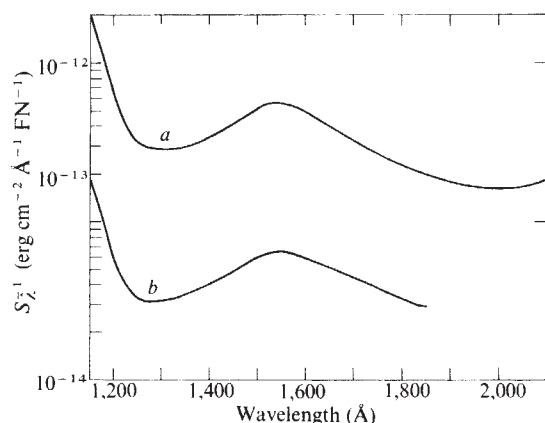


Fig. 8 Inverse absolute sensitivity curves for the short wavelength cameras in low dispersion. Multiplication of the indicated values by the linearised flux numbers obtained from the COMPARE data extraction program and division by the exposure time yields absolute fluxes for continuum sources in $\text{erg cm}^{-2} \text{s}^{-1} \text{\AA}^{-1}$. The curves are for the short wavelength redundant camera as used during the high priority observations and for the re-optimised short wavelength prime camera as it is now being used for routine observing. As the data reduction system scales fluxes differently for the two cameras, these curves should not be used directly for estimating exposure times. The real sensitivity difference between the prime and redundant cameras is only about a factor of two. *a*, SWR(4); *b*, SWP(3) reoptimised.

particularly in BD +75°325. The emission at 1,216 Å is typically present in long exposures at low dispersion and is due to resonant scattering of solar Ly α in the geocorona as discussed earlier. Note that in the quasar 3C273 the intrinsic Ly α and C IV 1,550 Å emissions are shifted longwards by the redshift $z = 0.158$.

Part of order number 106 from the short wavelength echelle spectra of four stars is shown in Fig. 7. Strong interstellar lines of O I 1,302.17 Å and Si II 1,304.37 Å are evident. Some variation in line position exists, because wavelengths are not corrected for velocities external to the spacecraft and because η UMa and HD153919 were observed in the large aperture, where pointing uncertainties cause wavelength errors. In η UMa, note the stellar features that are broad and shallow, in contrast to the sharp interstellar lines. The background used to find the net response in Fig. 7 was measured between the orders, where there is some residual due to lack of complete separation of the echelle orders at these wavelengths. Consequently, the cores of the interstellar lines tend to lie below zero. Improvements in the data processing system will be needed to correct this.

There are at least seven sources of noise in the IUE data: (1) quantum noise in the detected photoelectrons; (2) noise in the individual intensity transfer function (ITF) curves used to linearise the raw data values and reduce them to flux units; (3) inaccuracies in aligning the stellar images with the flat field images of the ITF sequences, so that each pixel uses the exact ITF curve intended; (4) random system noise in the tube readout and telemetry; (5) pseudo-periodic noise from other spacecraft subsystems; (6) noise in the pedestal level, which is the zero signal charge on the target after the camera preparation sequence, and (7) radiation background. The radiation background is discussed above and can be neglected for the short exposures described here. Care must be exercised to subtract in a smooth way any backgrounds present in the gross signal; otherwise, additional noise can be introduced by the data reduction. Table 5 lists some one- σ noise levels per sample point in IUE gross extracted spectra, where σ_{meas} includes noise from sources 1–6. The values of σ_0 are the quantum noise expected from the number of photoelectrons required to produce a net signal of 90% of saturation. For the spectrum of η UMa obtained with the long wavelength redundant camera, the measured noise approaches the theoretical value. In the other cameras, the measured noise is about twice the theoretical limit.

The low resolution data best define the photometric reproducibility of the IUE system, because the background can

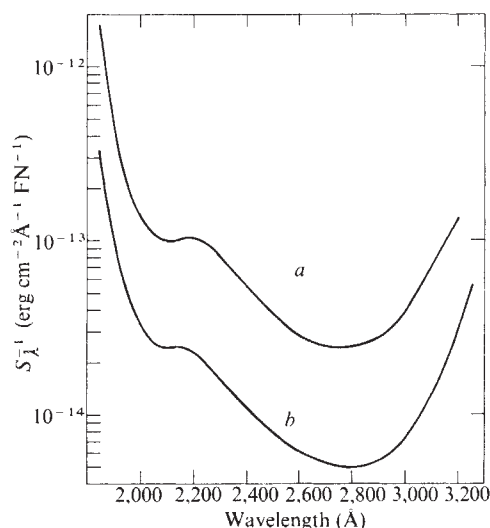
be measured reliably on either side of the spectrum. Studies of the photometric stability were therefore made by taking low dispersion spectra of the same star (HD60753, B3 IV, $V = 6.7$ mag) at periodic intervals in the large slot of each spectrograph. These data were obtained from 21 March–15 May 1978, and indicate a change in the sensitivity of the system. Over the full 55 days, the data showed decreases in sensitivity of 5% at 1,400 Å and 12% at 1,700 Å (taken with the short wavelength prime camera) and decreases of 5% at 2,150 Å and 9% at 2,750 Å (taken with the long wavelength redundant camera). The reasons for these changes are not yet understood and they will continue to be investigated as part of an overall programme to maintain a photometrically calibrated system.

Preliminary absolute sensitivity curves for low resolution IUE spectra have been obtained by observing hot stars with absolute energy distributions established from previous satellite and rocket observations. This procedure was preferred to a pre-launch calibration, as major efforts have been undertaken in the UV to calibrate existing satellite data¹⁰. The only extensive set of UV flux measurements available that covers the entire wavelength range spanned by IUE is the absolute OAO 2 spectrophotometry^{11,12}. Therefore, we have chosen to relate the preliminary IUE sensitivity curves to the absolute spectrophotometry on the OAO 2 system. Figures 8 and 9 show preliminary inverse sensitivity curves for all four cameras. The curves for the long wavelength redundant and short wavelength prime cameras apply to their re-optimised performance. The curves for the long wavelength prime and short wavelength redundant cameras are as they were used during the high priority phase of the scientific commissioning. The quantity plotted, S_{λ}^{-1} ($\text{erg cm}^{-2} \text{\AA}^{-1} \text{FN}^{-1}$), and the absolute flux, F_{λ} ($\text{erg cm}^{-2} \text{\AA}^{-1} \text{s}^{-1}$), are related by

$$F_{\lambda} = S_{\lambda}^{-1} \frac{(\text{FN})}{t}$$

where t is the exposure time in seconds and the IUE flux number (FN) is the linearised extracted net spectrum. Because of the details of the linearisation that produces FN, the sensitivity curves do not represent actual differences between cameras. The curves plotted are for large aperture exposures and a 10 pixel high spectral extraction with the program COMPARE. For small aperture exposures, where the light loss will depend on the telescope focus and the object positioning, these curves can only be used in the relative sense. The curves are based on short exposures (~ 0.6 – 2 s) of μ Columbae (HD38666, $V = 5.17$) timed with the aperture select mechanism in order to avoid uncertainties associated with the rise and fall times of the camera high voltage. The timing uncertainty associated with this

Fig. 9 As for Fig. 8 but for the long wavelength cameras. *a*, LWP(1); *b*, LWR(2) reoptimised.



mechanism may approach 10% for the shortest exposures. The sensitivity curves of Figs 8 and 9 imply short and long wavelength limits at about 1,150 and 3,200 Å, respectively. These sensitivity curves were obtained using preliminary intensity transfer curves to linearise the raw camera responses and a data extraction system that requires revisions. They should permit the transformation of IUE data into the OAO 2 spectrophotometric system with a relative accuracy of ~10%. However, the final result is also subject to the additional uncertainties of the OAO 2 absolute calibration, which seems to be accurate to about $\pm 5\%$ for $\lambda > 1,800$ Å and $\pm 20\%$ for $\lambda < 1,800$ Å. Figure 10 shows the result of using the derived short wavelength redundant camera calibration to compute a stellar flux distribution. The bump near 1,500 Å is located just where the OAO 2 flux distributions lie furthest above the measurements of the TD-1 satellite and other experiments¹⁰.

Because of its importance, a completely independent calibration, based on the same principles as above but using the TD-1 (S2/68) absolute fluxes as standard, was carried out during the commissioning phase. This is described by Boksenberg *et al.*⁶, and shows differences from that above (for example, it does not produce a bump near 1,500 Å) and further work is needed to establish the best sensitivity curves for general IUE use. In the work described in the following papers, both of these early calibrations have been used, that based on TD-1 for the interstellar⁴ and extragalactic⁶ work, and that based on OAO 2 for the rest^{2,3,5,7}.

Data reduction

Figure 11 shows the flow of the various types of images through the major steps of the data reduction system. Wavelength calibration lamp images are the emission line spectra from the on-board Pt-Ne source. Sequences of flat field images allow the nonlinear intensity transfer function (ITF) to be measured for each pixel in the image field. The standard star images are used to determine the absolute instrumental sensitivity by comparison with previous rocket and satellite spectra.

The first manipulation performed on each image is a geometrical transformation to make each pixel in each image

Fig. 10 *a*, The IUE spectral response for HZ43 produced by the image processing system after ITF correction. *b*, The absolute flux estimate for the object after application of the preliminary absolute calibration.

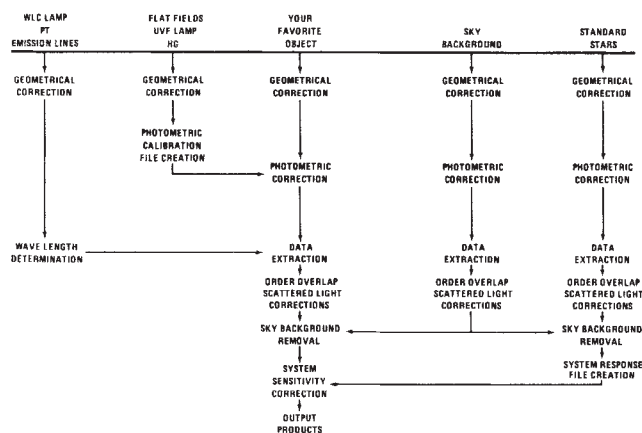
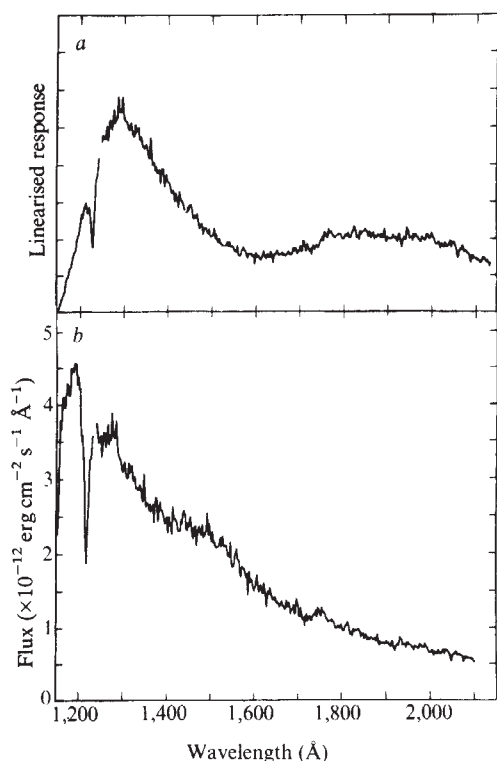


Fig. 11 The flow of images through the IUE spectral image processing system, showing the major image processing and data reduction tasks to be performed on each type of image.

refer to the same point in the focal plane at the camera faceplate. This removes the distortion inherent in the electron optics of the SEC tube, and hence the spectral orders are straightened out so that simpler fitting functions can be used to find the orders. However, in doing this transformation the data have been resampled, which results in some smoothing, each pixel in the transformed image being the linearly weighted average of the four adjacent pixels in the raw image. The next step is the photometric or ITF correction, which involves each point in the 768×768 pixel image being transformed with its own ITF curve. There are 12 levels of exposure in the ITF tables, with multiple images averaged together at each exposure level to reduce noise. Each pixel of the image being reduced is corrected by interpolation in its own 12-point table. The result is an image that is related linearly to the input flux on the camera, but has had no wavelength-dependent sensitivity and data corrections applied to it.

The next step is to extract the spectral data from the images. The location of the spectrum must be determined and this is obtained from known emission lines in the spectrum of the Pt-Ne calibration lamp, whose position, order number, and wavelength are fitted to the dispersion equations given above. When spectra are taken through the large entrance aperture, the wavelength scale derived from Pt-Ne calibration images obtained through the small aperture must be shifted linearly by about 30 pixels. In this case, the zero point of the resulting wavelengths may be slightly in error due to uncertainties in the exact centering of the target image within the 10×20 arc s aperture. After wavelengths have been assigned, the extraction of the spectrum is accomplished by passing one pseudo slit with a height of 5 diagonal pixels along the orders and another pseudo slit 1 diagonal pixel high between the orders. The inter-order background is smoothed and subtracted from the spectrum. This net spectrum must still be corrected for the system sensitivity. In high dispersion, a 'ripple correction' is applied to remove the effect of the echelle blaze. The data are then multiplied by the system sensitivity curves derived from standard stars.

Several problems remain in the data reduction as of May 1978, one of the major ones being the proper separation of the high dispersion orders. In the extraction of the high dispersion spectrum the smoothed interorder background is subtracted. This is adequate for the orders that are far enough apart for the inter-order background to be a true measure of the local background. However, as the order separation decreases at the shorter wavelengths, the inter-order background becomes an over-correction in both spectrographs. A major complication is that a variation of step size caused by the dynamic refocus voltage in the cameras makes the echelle orders deviate from straight lines, even in a geometrically corrected image. Therefore, in the refocus regions the pseudo slits that track the spectrum and the interorder background tend to lose or gain

signal, respectively. Thus, a small dip of typically 10% in an extracted spectrum becomes an even larger error when the anomalously high inter-order background is subtracted. Clearly, much more work is needed on this and other outstanding data reduction problems before an adequate image processing system is available for guest observers.

IUE observatory operations

The IUE observatory consists of the flight system plus the ground system. The flight system includes the spacecraft, the telescope, and the scientific instrumentation. The ground system includes the NASA IUE operations control centre at Greenbelt and the ESA's Villafranca satellite tracking station. The IUE observatory is designed to make maximum use of the continuous contact offered by the geosynchronous orbit. Guest observers are able to come to the observatory's science operations centres and actively participate in the real time control of their observations. In addition, the guest observer may take advantage of observing opportunities as they arise. Experienced resident astronomers and other observatory staff members assist the guest observer by providing advice on program planning, instrument operation, and data processing techniques. A guest observer usually arrives at the appropriate ground station a few days early to become familiar with the operations and data processing systems and to plan the final details of the observing program in consultation with resident astronomers.

The real time operations interface between the guest observer and the IUE is an interactive control and display terminal. This is manned by a telescope operator who is a specialist in spacecraft manoeuvring, target acquisition and instrument operation. The guest observer also sits at the display, and in consultation with the resident astronomer and telescope operator, directs critical aspects of the observation, including target acquisition, instrument operation, and data evaluation. The

interactive display provides the observer with all the information needed to plan slews, and also presents the field of view of the telescope to allow target identification. A quick-look spectrum is displayed as soon as the image is transmitted to the ground, and this allows the observer to evaluate the data quality and decide whether to repeat the observation or proceed to the next. The reduced data are available in the form of photographic reproductions of the raw and corrected images, verification plots of intensity versus wavelength, and magnetic data tapes. These materials are normally delivered to the observer 1 or 2 days after the observation.

Observers are given exclusive rights to their observations for 6 months after receipt of their data. The only exception to this rule is the right of an ESA observer to request full data on a duplicate target observed during a previous ESA shift. After 6 months, the data are deposited in the national space science data centre at Goddard Space Flight Center and in the data centres of the UK and the ESA. These data are available on request. Observers will be encouraged to use the data centres, rather than needlessly repeat observations; they may also find it advantageous to collaborate with other astronomers having similar observing programmes so that their combined observing time can be used to the greatest advantage.

1. Boggess, A. *et al.* *Nature* **275**, 356–361 (1978).
2. Heap, S. R. *et al.* *Nature* **275**, 369–372 (1978).
3. Linsky, J. L. *et al.* *Nature* **275**, 373–378 (1978).
4. Grewing, M. *et al.* *Nature* **275**, 378–384 (1978).
5. Dupree, A. K. *et al.* *Nature* **275**, 384–387 (1978).
6. Boksenberg, A. *et al.* *Nature* **275**, 388–398 (1978).
7. Lane, A. L. *et al.* *Nature* **275**, 398–399 (1978).
8. Mount, G. H., Yamasaki, G., Fowler, W. & Fastie, W. G. *Appl. Opt.* **16**, 591 (1977).
9. Allen, K. G. R., Anderson, B. E., Boksenberg, A. & Oliver, M. B. *Adv. Electron. Electron Phys.* **40A**, 223–237 (1976).
10. Strongylis, G. J. & Bohlin, R. C. NASA TM 78035 (1977).
11. Bless, R. C., Code, A. D. & Fairchild, E. T. *Astrophys. J.* **203**, 410 (1976).
12. Code, A. D. & Meade, M. *Wisc. Astrophys. No.* 30 (1976).

IUE observations of hot stars: HZ43, BD + 75°325, NGC6826, SS Cygni, η Carinae

**S. R. Heap, A. Boggess, A. Holm
D. A. Klinglesmith, W. Sparks, D. West
& C. C. Wu**

Goddard Space Flight Center, Greenbelt, Maryland 20771

A. Boksenberg, A. Willis & R. Wilson
University College London, Gower St, London WC1, UK

F. Macchetto
Astronomy Division, ESTEC, Noordwijk, Holland

P. L. Selvelli
ESA Satellite Tracking Station,
Villafranca Del Castillo, Madrid, Spain

D. Stickland
SRC Appleton Laboratory, Slough, Berks, UK

J. L. Greenstein
California Institute of Technology, Pasadena, California 91109

J. B. Hutchings
Dominion Astrophysical Observatory, Victoria, BC, Canada V8X 3X3

A. B. Underhill
Institut d'Astrophysique, Paris

R. Viotti
Laboratorio di Astrofisica Spaziale, Frascati, Italy

J. A. J. Whelan
Institute of Astronomy, Madingley Rd, Cambridge, UK

During the commissioning phase of IUE observations were made of a selection of hot stars: the white dwarf HZ43, the hot subdwarf BD + 75°325, the nucleus of the planetary nebula NGC6826, the dwarf nova SS Cygni, and the peculiar object η Carinae. The observations were made in the low dispersion mode and the data have been reduced using a preliminary calibration. The results are presented and discussed for each object.

THE IUE now provides the opportunity for major advances in the study of hot stars at UV wavelengths. The ability to detect faint sources now permits: (1) the study of hot, subluminal stars which as a group represent the avenues to the endpoints of evolution for the great majority of stars; (2) the study of hot, massive stars in external galaxies (the Magellanic Clouds); and (3) the study of hot stars (and their surroundings) which are heavily reddened by the interstellar medium. Another basis for this opportunity is the provision for complete spectral coverage from 1,150 to 3,200 Å, which permits: (1) a detailed study of a

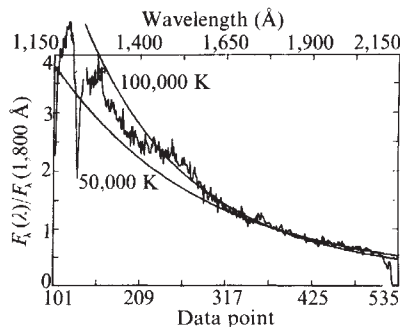


Fig. 1 Relative flux distribution of HZ43. Two abscissae are given: below, data point number of the spectrum extracted from the image; and above, the wavelength. The break in the spectrum near 1,230 Å is due to a reseau.

spectral region usually inaccessible to Copernicus, roughly 1,450–1,800 Å, which contains several lines (for example, C IV $\lambda\lambda$ 1548, 1550; He II λ 1640; N IV λ 1719) which are very useful diagnostics of the physical conditions in hot stars; and (2) the study of flux distributions of hot stars over a wavelength interval wide enough to distinguish differences. Finally, the imaging capability of IUE yields an efficiency and, potentially, a freedom from time-dependent errors associated with scanning spectrometers. This capability permits programmes of study based on repeated observations of variable hot stars, such as the Ap stars, binaries and pulsating stars.

During the commissioning period of IUE a sample of hot stars were chosen for observation. The main criterion for selection was that they should represent classes of hot stars deserving thorough study but, by virtue of their faintness, were previously inaccessible to UV spectroscopy. Because of the short time allotted for these observations it was decided to concentrate on the hot, subluminescent stars but to try to sample as many types of this class as possible. In addition, the object η Car was selected because of its peculiarity and heavy reddening. All spectra were obtained at low dispersion and have a nominal resolution of 6 Å. Table 1 lists the stars observed.

The emphasis of the work described here was to assess the usefulness of IUE low-dispersion spectra to conduct studies of hot stars. However, hot stars were observed at high dispersion during the commissioning period as parts of investigations concerning X-ray sources¹ and the interstellar medium².

Observations and reductions

Table 2 lists some of the observational data: the object, the image number, date of observation (day number of year 1978), the entrance aperture (small or large), the exposure time and comments about the quality of the spectrum. Most of the latter are based on an inspection of photographic reconstructions of the raw spectral data which allow qualitative assessments of such characteristics as general exposure level, system focus, systematic noise, and effects of particle radiation. The exposure levels of most spectra seem to be near optimum. In the short wavelength spectrograph, the small decline in system sensitivity towards

shorter wavelengths is compensated for by the characteristic increase in flux of the hot, unreddened stars under study here. Hence, one well-exposed spectrum is useful for the full spectral range, from $\sim$ 1,150 to 2,100 Å. The same generalisation holds true to a lesser extent for the long wavelength spectrograph covering the spectral region from 2,100 to 3,300 Å. The spectrograms generally show a low background level, and except for SWP 1141, which was obtained when the telescope was pointed near the Earth's limb, geocoronal Ly α is not present.

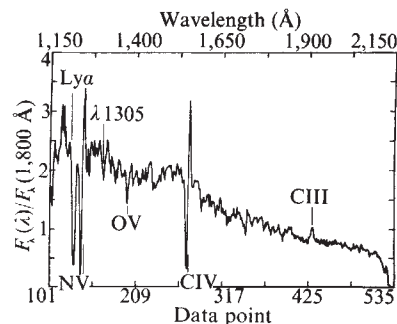


Fig. 3 Relative flux distribution of NGC6826.

The spectral images were reduced by a process, still under development, described in the preceding article³. Any error in the zero-point of the wavelength scale was then removed by assigning the laboratory wavelength to an identifiable feature, such as Ly α . This procedure was possible for all spectrograms except for those of η Car where no features were readily identifiable. The net spectrum was converted to absolute flux units with the preliminary system sensitivity functions derived from standard star observations (described in ref. 3). This sensitivity function tends to produce a swell near 1,500 Å and it is possible that this is ultimately due to the OAO 2 calibration, which formed the basis for this preliminary calibration. When applied to the spectral region below 1,200 Å where the spectral response is dropping rapidly, this early calibration is much less accurate and the data below Ly α in the following figures should be treated with caution.

Table 1 Hot stars observed by IUE

Object	Spectral type	V (mag)	Notes
BD +75°325	sdO	9.5	
HZ43	DAwk	12.8	EUV source
NGC6826	O3f	10.5	Central star of planetary nebula
SS Cyg	G + em. lines	9.1*	Dwarf nova
		11.0*	
η Car	pec	6.2	UV-IR object

* Magnitude estimated from FES counts at time of observation.

HZ43

The white dwarf HZ43 (spectral type DAwk) has long been known to be one of the hottest objects from ground-based observations^{4,5}. In recent years, it has been an object of considerable interest because it was the first extra-solar source of EUV radiation to be detected^{6,7}. The IUE spectrum of HZ43, shown in Fig. 1, shows no obvious spectral features other than Ly α . This is consistent with the interpretation that HZ43 is a white dwarf in which gravitational separation has already taken place so that hydrogen would be the only observable element in the stellar atmosphere. The Ly α feature observed here is probably essentially stellar rather than interstellar in origin, as the star is close to the Earth (about 60 pc)^{8,9} and the gravity of the star is high ($\log g \geq 7.7$) (ref. 10). Although Ly α is the only prominent feature, there appears to be 'fine structure' in the spectrum near 1,720 Å, which if confirmed could be ascribed to N IV.

Fig. 2 Relative flux distribution of BD +75°325.

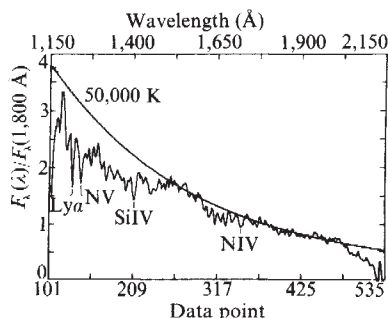


Table 2 Low dispersion spectral images obtained by IUE

Object	Image*	Date (day 1978)	Start (GMT)	Aperture	Exposure time (s)	Comments
BD + 75°325	SWR 1024	41	14.27	S	600	Camera noise
	LWP 1024	41	19.40	S	210	
	SWR 1057	57	07.51	L	1,200	
NGC6826	SWR 1058	57	11.50	S	360	Reoptimised camera parameters used
			12.14	L	180	
	LWR 1138	72	06.08	L	170	
SS Cyg	SWR 1050	57	15.51	S	150	Geocoronal Ly α present Reoptimised camera parameters used Zero-point of wavelength scale uncertain
			16.23	L	168	
	SWP 1141	72	22.42	L	480	
η Car	SWR 1060	57	00.16	L	150	Zero-point of wavelength scale uncertain
	LWR 1026	58	23.12	L	15	

* Images are designated by camera and sequence number: SWR = short-wavelength redundant camera; LWP = long-wavelength prime camera; SWP = short-wavelength prime camera; LWR = long-wavelength redundant camera.

The IUE flux distribution indicates a black body temperature between 50,000 and 100,000 K with negligible interstellar extinction, values that will be improved when the system is fully optimised and calibrated. Auer and Shipman¹⁰ from EUV and visual data estimate that the stellar parameters lie between the following limits: $T_{\text{eff}} = 55,000$ K, $\log g = 7.7$, $M = 0.4M_{\odot}$ and $T_{\text{eff}} = 70,000$ K, $\log g = 7.9$, $M = 0.5M_{\odot}$.

BD + 75°325

At a visual magnitude of 9.5, BD + 75°325 (spectral type sdO) is one of the brightest hot subdwarfs in the sky and has been studied in detail by Peterson¹¹. The only prominent lines in the visual spectrum are due to hydrogen, helium and nitrogen. The IUE spectrograms show a similar picture: absorption lines of He II, N IV and N V are strong, while lines of C III and C IV are below the threshold of detection. The long wavelength spectrum (not shown) indicates that the Paschen series of He II is stronger than in young O stars like ζ Pup¹², probably due both to helium enrichment and to Stark-broadening. The short-wavelength spectrum, shown in Fig. 2, indicates that nitrogen lines are strong, totally in absorption, and apparently at their rest wavelengths. The absorption in the spectral region, 1,600–1,750 Å, is real and most probably includes He II $\lambda 1640$. There is no evidence for the presence of the C IV resonance doublet at 1,550 Å. The apparent emission feature shortwards of Ly α is a noise spike.

In her study of the visual spectrum of BD + 75°325, Peterson¹¹ found that it was impossible to construct an LTE model atmosphere which predicts the observed characteristics of the star: the ionisation equilibrium between N III and N IV implies the atmospheric parameters, $T_{\text{eff}} = 55,000$ K, $\log g = 5.3$ and $N_{\text{He}}/N_{\text{tot}} = 0.6$, while the helium lines imply the parameters $T_{\text{eff}} = 45,000$ K, $\log g = 5.3$ and $N_{\text{He}}/N_{\text{tot}} = 0.6$. It is probable that much of the discrepancy is due to non-LTE effects. If so, it is

Fig. 4 Comparison of IUE spectrum of NGC6826 with Copernicus spectra of ζ Pup and o Per. The Copernicus spectra have been smoothed by a running box average whose width corresponds roughly to 6 Å.

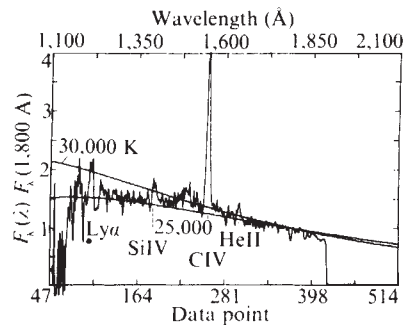
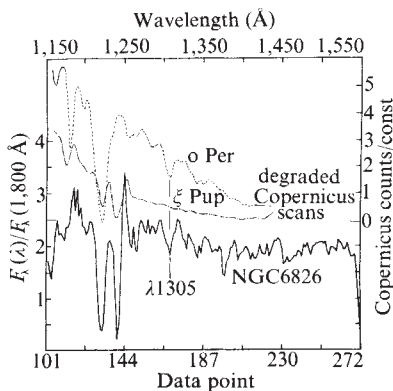


Fig. 5 Relative flux distribution of SS Cyg near quiescence.

important to obtain line strengths and profiles of important transitions of helium and nitrogen. High-dispersion spectra of BD + 75°325 could be obtained easily by IUE and should provide useful constraints on the physical properties and chemical abundances in the stellar atmosphere.

NGC6826

Figure 3 shows the short wavelength spectrum of NGC6826. The spectrum of the central star of this planetary nebula clearly has a high excitation, as would be expected from its visual spectral type of O3f (ref. 13). The most prominent lines are the resonance lines of N V $\lambda 1240$ and C IV $\lambda 1550$, as well as O V $\lambda 1370$ and N IV $\lambda 1718$ which arise from excited levels. The UV spectrum is very similar to that of the young star ζ Pup¹⁴, also an early Of star, with the strong lines all having P-Cygni profiles which indicate mass-loss from the central star. The interaction of the stellar wind with the surrounding nebula would be an interesting topic to pursue.

Although the stellar contribution is dominant, the spectrum of this object is a composite of stellar, nebular and interstellar spectra. Nebular emission of C III $\lambda 1909$ is clearly visible and fills the large aperture ($10'' \times 20''$). Nebular and/or interstellar absorption is also present: the strongest such feature, at 1,305 Å, is undoubtedly a blend of O I, Si II and P II resonance lines, all of which are of too low ionisation to belong to the stellar atmosphere. Figure 4 compares the spectral region containing this feature with degraded Copernicus U2 spectra of ζ Pup, which has very little reddening ($E(B - V) = 0.03$), and o Per, a B1 III star with heavy reddening ($E(B - V) = 0.32$). Note that the $\lambda 1305$ feature does not appear in ζ Pup but does appear in the spectrum of o Per with a strength comparable with that in the spectrum of NGC6826. If the relation between hydrogen column density and reddening given by Bohlin¹⁵ is used with the colour excess derived from Kaler¹⁶, a value of $\sim 10^{20} \text{ cm}^{-2}$ is obtained which is rather low for the observed line strength. There is therefore a possibility that some of the 1,305 Å feature may be nebular in origin but because of the loose nature of the correlation between hydrogen column density and reddening,

this would need to be established by other evidence. If the presence of nebula absorption lines were confirmed, for example, by high dispersion spectra, it would provide an important new tool for abundance analysis.

SS Cygni

The properties of the dwarf nova, SS Cyg, and dwarf novae in general have been reviewed by Warner¹⁷; since when, it was discovered¹⁸ and confirmed¹⁹ that SS Cyg is a soft X-ray source. The general explanation for these properties is that dwarf novae are binary systems in which material is being transferred from a low-mass secondary, through an accretion disk, to a white-dwarf companion. The 'hot spot', where the stream of transferred matter collides with the disk, and the disk as a whole are responsible for the UV and visual radiation.

UV spectra of SS Cyg were obtained on two occasions during the commissioning phase of IUE: the first during outburst, and the second near quiescence. The two spectra, which are shown in Figs 5 and 6, show the effect of an outburst. The most outstanding change is an increase in the continuous flux at outburst causing a reversal in the appearance of the C IV $\lambda 1550$ resonance doublet: near quiescence, it is an exceedingly strong emission feature, while during outburst it is a weak absorption or P-Cygni feature. In addition, emission at N V $\lambda 1240$, Si IV $\lambda 1400$ and He II $\lambda 1640$ are visible near quiescence, while during outburst these lines either disappear or go into absorption. The great strength of C IV emission near quiescence is important as an indicator of physical conditions in the disk.

A second change is an apparent increase in colour temperature during the outburst as expected from UV photometry of SS Cyg by Holm and Gallagher²⁰. The IUE spectra are similar to a black body temperature near 40,000 K during outburst maximum, and about 30,000 K near quiescence. As the interstellar extinction of this object is minimal ($E(B-V) \approx 0.08$) (ref. 20) these estimates should only slightly underestimate the intrinsic black-body colour temperatures of SS Cyg in this wavelength range. Of course, these values should be treated with caution because they are highly dependent on the absolute calibration, and they are not directly translatable to physical temperature until the mechanism of emission (probably optically thick free-free emission) and the range in physical conditions within the disk are specified.

η Car

One of the most remarkable objects in the sky is the star (or starlike object), η Car. Listed as a 4th mag star in Bayer's atlas of 1603, it brightened in the early 1800s until in 1843, it outshone every star in the sky except Sirius. For the past century, it has hovered between about 6th and 8th mag. The evolutionary status of this object is not understood, and theories about its nature²¹ are diverse, ranging from the hypothesis that it is a young, massive star either approaching the main sequence or somewhat evolved from it, to the hypothesis that it is a non-thermal or non-stellar object, such as a supernova remnant. Its present spectrum²²⁻²⁴ in the visual, consists of a continuum on which are superposed numerous emission lines (both permitted and forbidden) of Fe II, Cr II, Ni II, Na I, and so on, as well as the

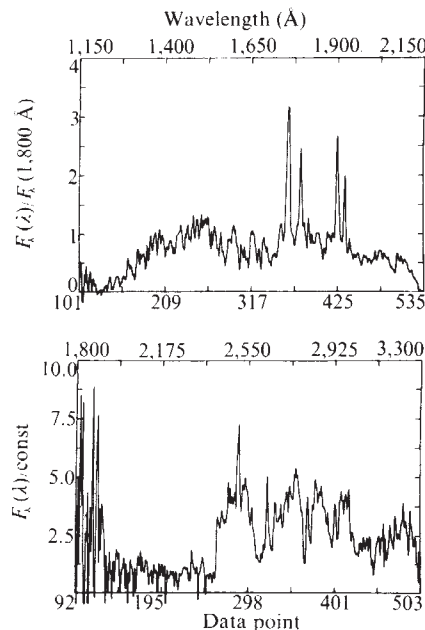


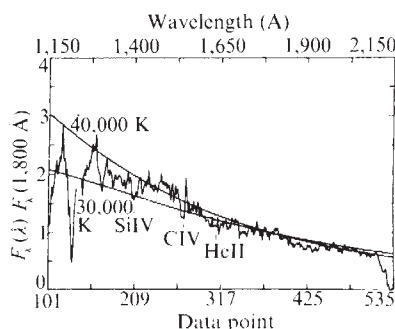
Fig. 7 Relative flux distributions of η Car.

hydrogen Balmer series. According to Rogers and Searle²⁴, the continuum contributes 70% and the emission lines 30% of the visual radiation. The object is heavily reddened both by interstellar material in the line of sight and by circumstellar gas and dust.

The IUE spectra of η Car, shown in Fig. 7, indicate that the UV spectrum is very complicated. One of the most outstanding characteristics is that the spectrum appears to change from an emission-line-dominated spectrum at longer wavelengths to an absorption-line-dominated spectrum at shorter wavelengths. According to Rogers and Searle, there are only a few spectral windows in the visual region of the spectrum which are free of emission lines, and in these regions the continuum is featureless. However, in the far-UV region from 1,150 to 1,600 Å, there appear to be about 30 absorption features visible, but it is not clear at this stage whether these are real or whether they are simply regions where there are no emission lines.

Another unusual characteristic of the UV spectrum is a flux discontinuity near 2,400 Å. Although interstellar extinction must play a part in depressing the flux in this spectral region, the drop is too precipitous to be ascribed totally to the interstellar $\lambda 2200$ extinction bump. Possibly it is due to a chance lack of emission lines below 2,400 Å. In any case, the discontinuity greatly increases the difficulty of inferring the amount of extinction from the observed UV flux distribution.

Fig. 6 Relative flux distribution of SS Cyg during outburst.



1. Dupree, A. *et al.* *Nature* **275**, 384-387 (1978).
2. Grewing, M. *et al.* *Nature* **275**, 378-384 (1978).
3. Boggess, A. *et al.* *Nature* **275**, 361-369 (1978).
4. Greenstein, J. L. in *Stellar Atmospheres* (ed. Greenstein, J. L.) 698 (University of Chicago Press, 1960).
5. Greenstein, J. L. & Eggen, O. *Astrophys. J.* **141**, 83 (1965).
6. Lampton, M., Margon, B., Paresce, F., Stern, F. & Bowyer, S. *Astrophys. J.* **203**, L71 (1976).
7. Hearn, D. R. *et al.* *Astrophys. J.* **203**, L21 (1976).
8. Margon, B. *et al.* *Astrophys. J.* **209**, 525 (1976).
9. Wagman, N. E. *Astr. J.* **72**, 957 (1967).
10. Auer, L. H. & Shipman, H. L. *Astrophys. J.* **211**, L103 (1977).
11. Peterson, A. V. thesis, Calif. Inst. Technol. (1970).
12. Morton, D. C. & Underhill, A. B. *Astrophys. J. Suppl.* **33**, 83 (1977).
13. Heap, S. R. *Astrophys. J.* **215**, 609 (1977).
14. Stecher, T. P. *Astrophys. J.* **159**, 543 (1969).
15. Bohlin, R. C. *Astrophys. J.* **200**, 402 (1974).
16. Kaler, J. *Astrophys. J. Suppl.* **31**, 517 (1976).
17. Warner, B. in *Structure and Evolution of Close Binary Systems* (eds Eggleton, P., Mitton, S. & Whelan, J.) 85-140 (Reidel, Dordrecht, 1976).
18. Rappaport, S., Cash, W., Doxley, R., McClintock, J. & Moore, J. *Astrophys. J.* **188**, L5 (1974).
19. Ricketts, M. J., King, A. R. & Raine, D. C. preprint (Univ. Leicester, 1978).
20. Holm, A. V. & Gallagher, J. S. *Astrophys. J.* **192**, 425 (1974).
21. Davidson, K. *Mon. Not. R. astr. Soc.* **154**, 415 (1971).
22. Thackeray, A. D. *Mon. Not. R. astr. Soc.* **113**, 211 (1953).
23. Gaviola, E. *Astrophys. J.* **118**, 234 (1953).
24. Rogers, A. W. & Searle, L. *Mon. Not. R. astr. Soc.* **135**, 99 (1967).

IUE observations of cool stars: α Aurigae, HR1099, λ Andromedae, and ε Eridani

J. L. Linsky, T. R. Ayres, G. S. Basri & N. D. Morrison

Joint Institute for Laboratory Astrophysics, NBS
and University of Colorado, Boulder, Colorado 80309

A. Boggess, F. H. Schiffer III & A. Holm

Goddard Space Flight Center, Greenbelt, Maryland 20771

A. Cassatella & A. Heck

ESA Satellite Tracking Station, Villafranca del Castillo, Madrid, Spain

F. Macchetto

Astronomy Division, ESTEC, Noordwijk, Holland

D. Stickland

SRC, Appleton Laboratory, Slough, Berks, UK

R. Wilson

University College London, Gower St, London WC1, UK

C. Blanco

Osservatorio Astrofisico, Catania, Italy

A. K. Dupree

Harvard-Smithsonian Center for Astrophysics, Cambridge,
Massachusetts 02138

C. Jordan

Department of Theoretical Physics, 1 Keble Rd, Oxford, UK

R. F. Wing

Astronomy Department, Ohio State University, Columbus, Ohio 43210

Initial IUE observations of four cool stars are reported. Observed fluxes and surface fluxes are given for several UV emission lines in the spectral range 1,175–2,000 Å, obtained at low and high dispersion with the short-wavelength spectrograph and camera. These lines are formed in the outer atmospheres of these stars, in regions presumably analogous to the solar chromosphere and transition region. The surface fluxes in the lines increase along the sequence: quiet Sun, ε Eri, λ And, α Aur, and HR1099. The 2.8-d RS CVn-type binary HR1099, observed on 1 March 1978 near the end of a major flaring episode, has line surface fluxes ~ 100 times that of the quiet Sun, similar to those seen in solar flares. Line profiles and flux ratios in multiplets for Capella are presented, and comments given on the opacity of the lines and on a tendency of line width to increase with temperature of formation.

THE IUE has made it feasible to study for the first time UV spectra of many late-type stars. These emission-line spectra originate in chromospheres, coronae, and circumstellar envelopes, and they contain information about mass loss from late-type stars. Before IUE, Copernicus and rocket- and balloon-borne experiments could study only the strongest lines in a few very bright stars. Here we present observations of α Aur (Capella), λ And, HR1099, and ε Eri. These data were obtained during the science commissioning phase of IUE and Table 1 gives the circumstances of the observations. Although these data were obtained before reoptimisation of the telescope and detectors, they indicate the ultimate capabilities of IUE. In this first sampling of cool stars, we have chosen those expected to have bright emission-line spectra and have concentrated on the 1,175–2,000 Å region, where most of the important emission lines are located.

α Aurigae (Capella)

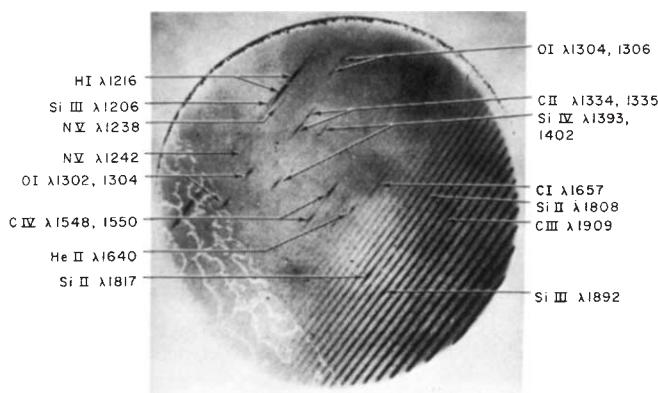
Capella is a double-line spectroscopic binary system, consisting of a G5 III primary (Capella A) and a G0 III secondary (Capella B)¹. It was chosen as the first IUE scientific target because it is easily identifiable and is known to have several bright UV emission lines.

Evidence for a chromosphere in at least one component consists of emission cores in the Ca II H and K lines², the Mg II h

(λ 2803) and k (λ 2796) lines³, and H I Ly α ^{4,5}. Using the Ca II and Mg II line fluxes and identifying the chromospheric spectrum with Capella A^{3,4}, Kelch *et al.*² derived a model for the chromosphere and upper portion of the photosphere of this star. Material at temperatures above 20,000 K is indicated by UV emission lines of C II–IV, Si II–IV, N V, and O VI^{4–6}. Haisch and Linsky⁷ have interpreted these data in terms of a geometrically narrow conductively heated transition region similar to the transition region in a solar plage. Even hotter material is indicated by X-ray observations^{8,9} and possibly by the observed strong He I λ 10830 absorption¹⁰. The very recent HEAO-1 observations of Cash *et al.*¹¹ are consistent with a 10^7 K isothermal plasma with solar composition.

In the two exposures obtained of Capella on 9 February 1978, the large aperture ($10'' \times 20''$) included nearly all of the slightly out-of-focus stellar image. The first image was a high-dispersion exposure of the long-wavelength spectrum (1,900–3,200 Å) with the prime camera (image LWP 1022). This spectrum shows the Mg II λ 2796, 2803 lines in emission and the photospheric absorption line spectrum well exposed for 2,530–2,880 Å and 3,100–3,270 Å, but largely overexposed for 2,880–3,100 Å. The second observation was a high-dispersion exposure of the short-wavelength spectrum (1,175–2,100 Å) with the redundant camera (image SWR 1022). Qualitatively this consists of emission lines shortwards of 1,600 Å, and a photospheric absorption-line spectrum with superimposed emission lines at longer wavelengths. This spectrum is shown in Fig. 1 with the

Fig. 1 Short-wavelength high-dispersion raw spectrum of Capella (image SWR 1022) in the IUE echelle format. Important emission lines are identified.



important emission lines identified. Table 2 lists measured wavelengths of emission lines or blends and the spectrum orders. Also included are proposed identifications giving ion, multiplet number, and laboratory wavelength based on the solar line lists of Burton and Ridgeley¹² and Doschek *et al.*¹³ The relative wavelength scale was obtained from a previous Pt-Ne lamp exposure, but with 140.3 km s⁻¹ added to fit optimally the measured wavelengths of the 13 bright single narrow emission lines. This correction allows for errors in the grating dispersion constants and for the stellar radial velocity and heliocentric motions, but it prevents a velocity measurement of the emission line spectrum relative to the mean velocity of the binary system. The measured wavelengths of all bright single identified lines agree with the laboratory wavelengths to within 0.05 Å.

Relative line fluxes were determined by subtracting the background and the stellar continuum if present, and then integrating the net signal across the emission lines. These fluxes, in IUE flux units, are given in Table 2 for α Aur and for a preceding 45-s exposure of the standard star η UMa obtained with similar observing parameters. Where stellar or interstellar absorption lines are present in the η UMa spectrum, we interpolated between continuum regions on either side. Dividing the two spectra provides the best means of estimating fluxes at Earth for Capella, corrected for the echelle blaze characteristics and variable detector sensitivity with position. We have assumed the η UMa absolute fluxes of Code and Meade¹⁴, again interpolating between absorption features where necessary shortwards of 1,400 Å.

Table 1 Summary of observations

Star	Spectral type	m_V	Binary period (d)	Phase (d)	Image no.	Dispersion	Exposure midpoint (JD 244+)	Exposure (min)	Comments
α Aur (Capella)	G5 III + GO III	0.09	104.023	77.670 77.791	LWP 1022 SWR 1022	High High	3549.321 3549.442	3 70	
λ And	G8 III-IV	3.88	20.5212	13.438 13.715	SWR 1044 SWR 1045	High Low	3560.146 3560.423	240 152	Underexposed
HR1099	G5 V	6.0	2.83782	2.4488	SWR 1065	Low	3568.783	100	
ϵ Eri	K2 V	3.73	—	—	SWR 1067 SWR 1069	Low Low	3569.383 3569.619	38 140	Underexposed

Table 2 Emission lines observed in Capella

Measured wavelength (Å)	Spectral order	Identification ion	FWHM (Å)	IUE flux		Flux at Earth (erg cm ⁻² s ⁻¹)	Surface flux	Surface flux (α Aur/quiet Sun)	Comments
				α Aur	η UMa*				
1,175.26	117	C III(4)	1,175.1	8.0(2)§	7.2(4)§				
1,206.45§	114	Si III(2)	1,206.51	1.5(4)§	8.0(4)§	2.0(-11)§	6.6(4)§	20§	Very uncertain Uncertain due to Ly α spillover from adjacent order
1,215.58§	113	H I(1)	1,215.66	1.51	2.9(5)				
	113	O V	1,218.39		—				Not apparent
1,237.28§	111	?		5.4(3)§	2.7(5)	2.9(-12)§	9.4(3)§		Uncertain
1,238.70§	111	N V(1)	1,238.82	0.75	1.0(4)§	2.9(5)	5.0(-12)§	32§	
	111	[Fe XII]	1,242.03		—				Not apparent
1,242.80	111	N V(1)	1,242.80	0.67	3.8(3)§	4.4(5)	1.3(-12)§	12§	
1,302.14	106	O I(2)	1,302.17	0.50	1.6(4)	4.2(5)	6.5(-12)	17	
1,304.83	106	O I(2)	1,304.86	0.55	1.6(4)	3.2(5)	8.2(-12)		
1,304.85	105	O I(2)	1,304.96	0.48	1.1(4)	1.8(5)	1.0(-11)		
1,305.98	106	O I(2)	1,306.03	0.58§	1.2(4)§	2.7(5)	7.7(-12)§	20	Near edge of order Obscured by blemish
1,306.05	105	O I(2)	1,306.03	0.41	1.3(4)	2.5(5)	8.6(-12)		
1,334.55	103	C II(1)	1,334.53	0.85	3.2(4)	4.0(5)	1.3(-11)		
1,335.69	103	C II(1)	1,335.66	0.78	4.0(4)	4.4(5)	1.5(-11)	20	
	103	[Fe XII]	1,349.38						Not apparent
1,355.50	102	O I(1)	1,355.60	0.36§	3.4(3)§	3.2(5)	1.7(-12)§	17§	
	102	O I(1)	1,358.51						Present but weak
1,393.75	99	Si IV(1)	1,393.76	0.63	2.2(4)	3.5(5)	9.2(-12)	17	
1,401.37§	98	O IV	1,401.16	0.96§	1.3(4)§	2.9(5)	6.4(-12)§	160§	
1,402.76	98	Si IV(1)	1,402.77	0.71	2.0(4)	3.1(5)	9.5(-12)	47	
1,404.6§	98	O IV	1,404.81			3.5(5)			Possibly present
1,548.20	89	C IV(1)	1,548.20	1.02	4.1(4)	1.9(5)	2.8(-11)	25	
1,550.80	89	C IV(1)	1,550.77	0.75	2.0(4)	2.0(5)	1.3(-11)	21	
1,560.88	88	C I(3)	1,560.6†	0.24§	7.6(2)	1.4(5)	7.0(-13)		
1,561.38	88	C I(3)	1,561.4†	0.38§	2.3(3)	1.5(5)	2.0(-12)	4.5	
1,563.52	88	Fe II(45)	1,563.79			1.7(5)			
1,640.47	84	He II(12)	1,640.4†	0.53	1.5(4)	2.0(5)	7.9(-12)	20	Possibly present Additional features at 1639.93 (possibly due to Fe II 1640.15) and 1640.98
1,657.15§	83	C I(2)	1,657.1†	2.17	2.3(4)	2.1(5)	1.1(-11)	7.0	Peaks at 1656.33 (due to C I 1656.27) and at 1657.80 (due to C I 1657.91)
1,808.05	76	Si II(1)	1,808.01	0.42	3.3(4)	3.5(5)	8.1(-12)	4.0	
1,816.93	76	Si II(1)	1,816.93	0.43					
1,817.3§	76	Si II(1)	1,817.45		1.0(5)	3.5(5)	2.5(-11)	8.3	
1,892.00	73	Si III(1)	1,892.03	0.69	8.4(4)	3.3(5)	1.9(-11)		
1,908.71	72	C III	1,908.73	0.54	3.5(4)	3.9(5)	6.5(-12)		
1,992.88	69	?		0.32	1.1(4)	3.9(5)	1.7(-12)		
1,993.59	69	C I(32)	1,993.62	0.28	1.6(4)	4.1(5)	2.4(-12)		

* Per Å interpolated to line centre if necessary to avoid absorption lines.

† Multiplet, wavelength given is for the mean of the solar emission feature.

‡ Line not in emission in spatially-averaged solar spectrum.

§ Uncertain.

For stars with a bright continuum like η UMa, inter-order light is not a good measurement of the background because it includes the spillover of light from adjacent orders, especially in the short-wavelength portion of the spectral format, where the orders are most closely spaced. We have therefore determined the background for η UMa as a linear function of position perpendicular to the echelle dispersion direction such that the flux in the centre of the interstellar Ly α line is zero and the background equals the inter-order signal between the two longest-wavelength orders, which are the most widely separated.

Surface fluxes (fluxes per unit area at the stellar surface) for Capella are derived from the measured fluxes at Earth and an angular diameter for Capella A of 0.0072" (ref. 2). These surface fluxes were then divided by the corresponding quiet Sun disk average values of Rottman (personal communication), which were obtained near solar minimum.

Inspection of unsmoothed line profiles is a good way of assessing the quality of the IUE observations of cool stars. The Ly α line shown in Fig. 2 is very well exposed, with signal-to-noise ratio ~ 30 near the peak. However, the interstellar hydrogen absorption feature does not reach zero intensity at line

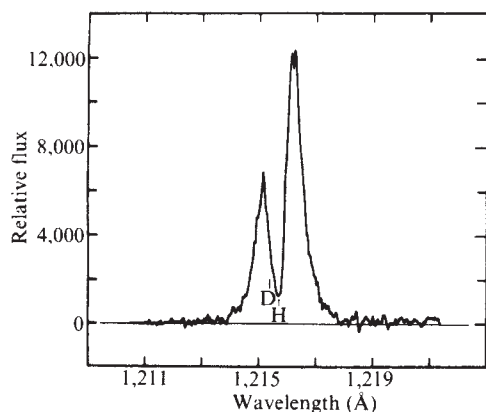


Fig. 2 Reduced profile of the Capella Ly α line in relative IUE flux units. Location of interstellar features due to hydrogen and deuterium Ly α is noted.

centre and interstellar deuterium absorption barely appears, indicating a degraded spectral resolution due to telescope mis-focus at the time of this early observation.

Weaker lines permit measurement of line fluxes and widths, but not details of the line profiles. The C IV lines in Fig. 3 have signal-to-noise ratio ~ 8 near their peaks and an integrated line flux ratio of 2.1, compared to the solar ratio of 1.8. The O I resonance lines shown in Fig. 4 have a similar signal-to-noise ratio to the C IV lines, except for the $\lambda 1306$ line which is located near the end of order 106. Finally the C II lines shown in Fig. 5 have peak signal-to-noise of perhaps 6, yet the line fluxes and line flux ratio can be fairly accurately measured as there are about 220 pixels per line.

ϵ Eridani

This bright K2 V star was observed as a prototype late-type dwarf and also as a single star against which to compare the binary systems Capella, λ And, and HR1099. Emission in the cores of the Ca II and Mg II resonance lines had earlier provided evidence for a chromosphere in this star. From measured surface fluxes 2.3 and 1.6 times those in the average quiet Sun for Ca II and Mg II, respectively, Kelch¹⁵ derived a chromospheric model.

We obtained two low-dispersion short-wavelength observations (see Table 1), but here discuss only the well-exposed SWR 1069 image. This spectrum was calibrated using the preliminary sensitivity curves for this camera¹⁶, and is compared with the quiet Sun spectrum of Rottman (personal communication) degraded to the low-dispersion resolution in Fig. 6.

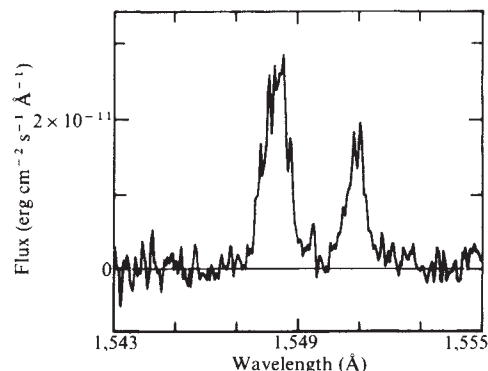


Fig. 3 Reduced profiles of the Capella C IV $\lambda\lambda 1548, 1550$ lines in absolute flux units.

Except for the relative weakness of the continuum at long wavelengths, which is expected for the cooler star, the spectrum of ϵ Eri is qualitatively similar to that of the Sun. The broad, saturated feature which covers the stellar Ly α and Si III $\lambda 1206$ lines is due to geocoronal Ly α emission filling the $10'' \times 20''$ large aperture.

Table 3 lists the emission features that clearly show up above the noise, their identification, and the fluxes at Earth. Only those features which appear in two or more stars are included. For analytical purposes, the most interesting quantities are the surface fluxes, which we determine from an angular diameter of 0.00229" derived by the $V-R$ relation of Barnes and Evans¹⁷. The surface fluxes for ϵ Eri and for the average quiet Sun are listed in Table 3.

λ Andromedae

λ And is a single-lined spectroscopic binary with a period of 20.5 d and is a long-period RS CVn-type binary variable. This star shows bright and variable emission cores in the Ca II (ref. 18), Mg II, and Ly α lines which may be a consequence of the primary filling a significant fraction of its Roche lobe¹⁹. We obtained both high- and low-dispersion short-wavelength images of λ And. The high-dispersion image is underexposed, but it weakly exhibits emission in the following lines: Ly α ; O I $\lambda\lambda 1302, 1304, 1306$; C II $\lambda\lambda 1334, 1335$; C IV $\lambda\lambda 1548, 1550$; He II $\lambda 1640$; Si II $\lambda\lambda 1808, 1817$; and Si III $\lambda 1892$. As a result of the faintness of this exposure, we observed the star again but in low dispersion. This spectrum was calibrated in a manner similar to that of ϵ Eri and is shown in Fig. 6. Table 3 gives line fluxes and surface fluxes, for which an angular diameter of 0.00274" was assumed.

The low-dispersion spectrum was obtained at spectroscopic phase 13.751 d with respect to the orbital elements of Walker²⁰. λ And has a photometric period of 50–55 d. In unpublished UVB photometry Hall found that λ And was brightest on JD 2,443,360 and 2,443,420. Thus our observations were obtained very near photometric maximum, when IUE was observing the unspotted and presumably least active hemisphere.

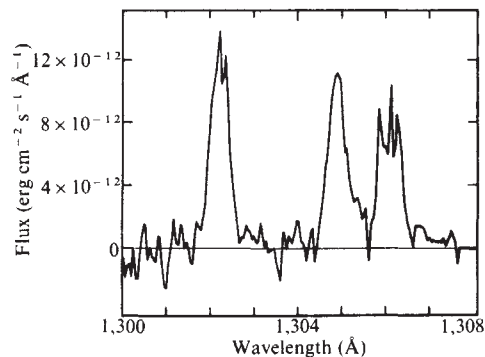


Fig. 4 Reduced profiles of the Capella O I $\lambda\lambda 1302, 1304, 1306$ lines in absolute flux units.

Table 3 Emission lines observed at low dispersion

Wavelength (Å)	Identification	Flux at Earth (erg cm ⁻² s ⁻¹)			Surface flux (erg cm ⁻² s ⁻¹)			
		HR1099	λ And	ϵ Eri	HR1099	λ And	ϵ Eri	Sun
1,175	C III(4)	4.4(-13)	3.0(-13)	—	1.8(5)	8.4(3)	—	1.6(3)
1,215	H I	SAT	SAT	SAT	—	—	—	2.1(5)
1,240	N V(1)+[Fe XII]?	1.4(-13)	1.9(-13)	1.4(-13)*	5.7(4)	5.3(3)	4.5(3)*	8.5(2)
1,263	Si II(4)	4.3(-14)*	2.5(-14)*	2.1(-13)*	1.8(4)*	7.0(2)*	6.8(3)*	1.1(3)
1,276	C I b1	2.4(-14)*	3.2(-14)*	9.9(-14)*	9.8(3)*	8.9(2)*	3.2(3)*	5.2(2)
1,281	C I(5)	6.4(-14)	—	1.2(-13)*	2.6(4)	—	3.9(3)*	—
1,304	O I(2)	6.0(-13)	2.1(-12)	4.6(-13)	2.5(5)	5.9(4)	1.5(4)	4.0(3)
1,335	C II(1)	9.0(-13)	6.7(-13)	4.7(-13)	3.7(5)	1.9(4)	1.5(4)	4.6(3)
1,357	O I(1)	6.7(-14)	2.5(-13)	—	2.7(4)	7.0(3)	—	3.4(2)
1,371	O V(7)	4.0(-14)	3.1(-14)*	—	1.6(4)	8.6(2)*	—	—
1,382	Fe II+Al III	2.1(-14)*	3.2(-14)*	—	8.6(3)*	8.9(2)*	—	—
1,394	Si IV(1)	2.3(-13)	3.2(-13)	1.5(-13)	9.4(4)	8.9(3)	4.9(3)	1.7(3)
1,403	Si IV(1)+O IV	2.3(-13)	5.0(-13)	1.8(-13)	9.4(4)	1.4(4)	5.8(3)	7.8(2)
1,425	Fe II(47)+S I(5)	2.4(-14)*	1.2(-13)	1.3(-13)*	9.8(3)*	3.3(3)	4.2(3)*	—
1,433	C I(65)+S I(5)	6.0(-14)	3.1(-14)*	9.0(-14)*	2.5(4)	8.6(2)*	2.9(3)*	2.1(2)
1,470	C I(36)+S I(4)	5.8(-14)	8.8(-14)	1.7(-13)*	2.4(4)	2.5(3)	5.5(3)*	5.2(2)
1,482	S I(4)+S I(3)	6.3(-14)	—	1.5(-13)*	2.6(4)	—	4.9(3)*	5.3(2)
1,530	Si II(2)	1.5(-13)	4.7(-14)	1.9(-13)*	6.2(4)	1.3(3)	6.2(3)*	2.3(2)
1,549	C IV(1)	1.3(-12)	1.8(-12)	5.3(-13)	5.3(5)	5.0(4)	1.7(4)	5.8(3)
1,561	C I(3)+Fe II(45)	1.4(-13)	—	3.7(-13)*	5.7(4)	—	1.2(4)*	2.0(3)
1,640	He II(12)+Fe II(43)	6.0(-13)	1.0(-12)	4.9(-13)	2.5(5)	2.8(4)	1.6(4)	1.3(3)
1,657	C I(2)+Fe II	3.6(-13)	5.9(-13)	3.6(-13)	1.5(5)	1.6(4)	1.2(4)	5.3(3)
1,670	Al II(2)+O III+Fe II	8.4(-14)*	3.3(-13)	8.8(-14)*	3.4(4)*	9.2(3)	2.9(3)*	1.5(3)
1,686	Fe II(40)	4.4(-14)	5.2(-14)*	6.7(-14)*	1.8(4)	1.5(3)*	2.2(2)*	—
1,696	Fe II(38)	7.7(-14)	4.5(-14)	1.7(-13)	3.2(4)	1.3(3)	5.5(3)	—
1,714	Fe II b1	5.2(-14)	5.9(-14)	2.9(-14)*	2.1(4)	1.6(3)	9.4(2)*	—
1,725	Fe II b1	5.2(-14)	4.9(-14)*	8.7(-14)*	2.1(4)	1.4(3)*	2.8(3)*	—
1,748	Ni II(5)	6.9(-14)	8.1(-14)	—	2.8(4)	2.3(3)	—	—
1,763	Fe II(101)+C I	—	2.8(-14)*	8.8(-14)*	—	7.8(2)*	2.9(3)*	—
1,788	Fe II	4.7(-14)	8.5(-14)	—	1.9(4)	2.4(3)	—	—
1,808	Si II(1)	2.1(-13)	7.1(-13)	3.3(-13)	8.6(4)	2.0(4)	1.1(4)	6.1(3)
1,817	Si II(1)	5.1(-13)	SAT	9.4(-13)	2.1(5)	—	3.1(4)	1.0(4)
1,842	Fe II(65)	2.8(-14)	3.2(-14)	—	1.1(4)	8.9(2)	—	—
1,855	Al III(1)+Fe III(63)	8.1(-14)	9.7(-14)	—	3.3(4)	2.7(3)	—	—
1,863	Al III(1)+Fe II b1	7.4(-14)	1.1(-13)	—	3.0(4)	3.1(3)	—	—
1,892	Si III(1)	1.5(-13)	2.5(-13)	7.2(-14)	6.2(4)	7.0(3)	2.3(3)	—
1,900	S I(1)+Fe	8.9(-14)	1.9(-13)	1.2(-13)	3.7(4)	5.3(3)	3.9(3)	—
1,909	C III	1.7(-13)	1.3(-13)	—	7.0(4)	3.6(3)	—	—
1,931	C I(33)	7.2(-14)	7.5(-14)	9.6(-14)	3.0(4)	2.1(3)	3.1(3)	—
1,948	Fe II(123)	8.4(-14)*	2.8(-14)	1.2(-13)	3.4(4)	7.8(2)	3.9(3)	—
1,993	C I(32)+Fe III(81)	2.4(-14)	1.7(-13)	6.6(-14)	9.8(3)	4.7(3)	2.1(3)	—

* Uncertain.

HR1099

Bopp and Fekel²¹ pointed out that HR1099 is a double-lined spectroscopic binary, and is one of the shortest-period RS CVn-type binaries. They show that the Ca II H and K and H α emissions are due mainly to the primary star. HR1099 also shows Ly α , Mg II h and k, radio, and X-ray emission²²⁻²⁴. We observed this object on 1 March 1978, towards the end of the most intense flaring episode yet observed for this system²⁵. The observed spectrum of HR1099 is shown in Fig. 6 and fluxes are given in Table 3. From the G5 V spectral type²⁶, we derived $V-R = 0.61$ and hence an angular diameter of $0.000644''$. Our observation occurred 2.4488 d after conjunction²⁷, when the primary, and presumably more active, component was near maximum velocity of approach²¹.

Discussion

Detailed analysis and atmospheric modelling of these data will be reported elsewhere, and both deeper exposures and observations at different phases of the binary targets would be helpful. In the meantime it is important to discuss and assess these new data qualitatively.

First, there is no clear evidence of 10^6 K or hotter material in the 1,175–2,000 Å spectral region, but as only very weak coronal lines such as [Fe XII] $\lambda\lambda 1242.03, 1349.38$ occur in this spectral range, this conclusion is not unexpected. These lines do not appear in the quiet Sun disk spectrum but they do appear off the limb above active regions and during flares²⁸. The hottest plasma clearly identified in our data has a temperature of 2.6×10^5 K, as indicated by the O V $\lambda 1371$ line in HR1099 and λ And. However, we point out that the broad feature at 1,355 Å in HR1099 and ascribed to O I (1) $\lambda 1355.6$ and $\lambda 1358.5$ could

include a contribution due to [Fe XXI] $\lambda 1354.1$, which is bright in solar flares²⁹.

Second, there are important qualitative differences among the five spectra shown in this paper (Figs 1 and 6). In both the Sun and Capella the photospheric absorption-line spectrum is clearly present at wavelengths longwards of 1,500–1,600 Å, but is very weak in the cooler stars. It is likely that Capella B is mainly responsible for the observed photospheric spectrum of Capella.

Line ratios often vary strongly from star to star. To facilitate comparison, Table 4 lists the ratios of accurately measured line surface fluxes to corresponding values for the average quiet Sun. Average values are given for ions whose lines are formed in similar temperature ranges³⁰, and we have included Ly α , Ca II, and Mg II data from other sources.

There are several trends in these data. First, the O I resonance lines are often much brighter than the other chromospheric ($T < 10^4$ K) lines, and the He II $\lambda 1640$ line is often much brighter than the other transition-region ($T > 3 \times 10^4$ K) lines. The O I anomaly is probably due to the Ly β pumping mechanism that operates in α Boo³¹ and the Sun³², and the He II $\lambda 1640$ line can be formed by recombination following photoionisation of He II by $\lambda < 227$ Å coronal photons³³. Thus, when the O I and He II lines seem anomalously bright, one may infer that these fluorescent processes are more important compared to collisional excitation than in the Sun.

In ϵ Eri the chromospheric and transition-region lines, except for He II $\lambda 1640$, are all about three times as bright as in the Sun; hence, the basic structure of the outer atmosphere of this prototype K dwarf bears similarities to that in the Sun. The factor of 3 increase in brightness of the chromospheric lines has been explained by a steeper temperature gradient in the lower

chromosphere¹⁵, which results in higher densities and thus more collisional excitation of the lines. If the transition region in ϵ Eri is geometrically thin and conductively heated as in the Sun, then the strength of its lines is proportional to the transition region pressure⁷ and this pressure should be a factor of 3 larger in ϵ Eri than in the quiet Sun. λ And is similar to ϵ Eri except that there seems to be twice as much material in the transition region.

Table 4 Line surface fluxes/average quiet Sun

Ions	Temperature	ϵ Eri	λ And	α Aur	HR1099	Ref.
C I, Si I, Si II	$4-6.5 \times 10^3$	2.4	4.2	4.9	70	This work
O I $\lambda\lambda 1302, 4, 6$	$\sim 6 \times 10^3$	3.8	16	19	63	This work
Ca II, Mg II	$\sim 6 \times 10^3$	2.0	4.6	3.7	50	2, 15, 22
Ly α	2×10^4	1.3*	3.9*	1.5*	50*	5, 22, 26
C II, C III	$4.5-9 \times 10^4$	3.3	4.7	20	100	This work
He II	$\sim 8 \times 10^4$	12	22	20	190	This work
C IV, Si IV, N V	$0.8-3 \times 10^5$	4.4	9.5	26	84	This work

* Uncorrected for interstellar absorption.

Capella and HR1099, by comparison, are very different from the quiet Sun. Capella shows considerably more material in its transition region than the Sun, whereas HR1099 shows large enhancements in both the transition region and the chromosphere. Since the surface fluxes for HR1099 are enhanced over the quiet Sun fluxes by a factor of 10^2 , similar to the typical enhancement of permitted lines during solar flares³⁴, a solar-like flare may have covered most of the stellar surface or a far more intense flare may have covered a small portion of the surface during this episode.

Third, for Capella the line surface fluxes for ions formed at $T > 3 \times 10^4$ K are typically 20 times larger than for the quiet Sun. By comparison, Vitz *et al.*⁵ obtained values only 10 times larger than for the quiet Sun. This apparent factor of two discrepancy is due mainly to the different angular diameters for Capella A assumed here (0.0072") and by Vitz *et al.*⁵ (0.0094") in converting observed flux to surface flux. Measured fluxes at Earth are in good agreement (within 15%) for the $\lambda\lambda 1206$, 1240, 1305, and 1550 features, but the IUE values are about 65% larger for the $\lambda\lambda 1335$ and 1397 features formed by the C II and the Si IV and O IV lines. We cannot say with certainty at this stage whether the discrepancy in the $\lambda\lambda 1335$ and 1397 fluxes is due to measurement error or represents real changes in the emission from Capella. We suspect, however, that the discrepancy results from the preliminary nature of our calibration scheme, as the $\lambda 1550$ fluxes, due mainly to C IV, agree and there is no evidence for large changes in emission-line fluxes with phase except for O VI.

Fourth, in the high-resolution Capella spectrum there is no systematic trend in wavelength residual with ionisation potential or presumed temperature of maximum abundance. Thus we see no evidence for the onset of a stellar wind at temperatures of 100,000–200,000 K, as previously seen in the O VI $\lambda 1032$ line (formed at 300,000 K) by Dupree⁴ at phase 38.2 d.

Fig. 5 Reduced profiles of the Capella C II $\lambda\lambda 1334, 1335$ lines in absolute flux units.

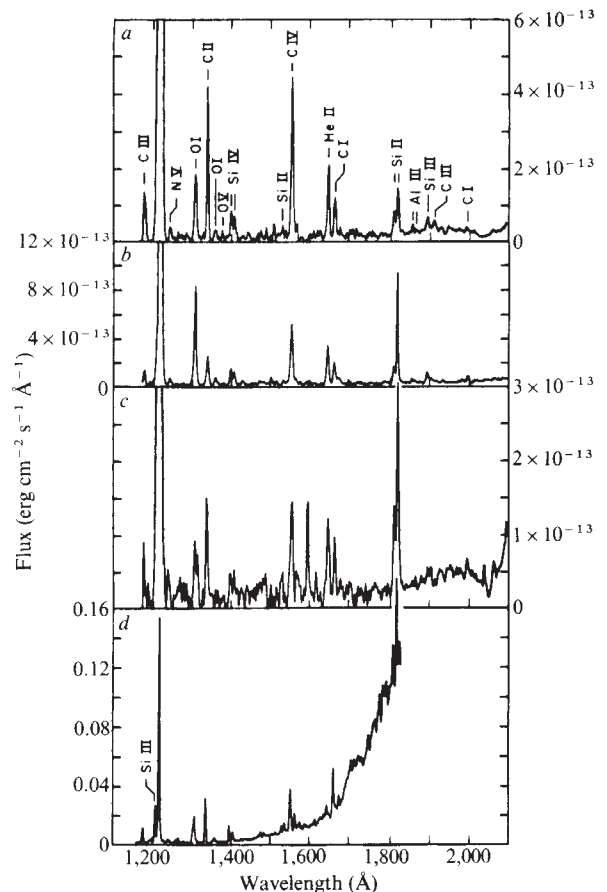
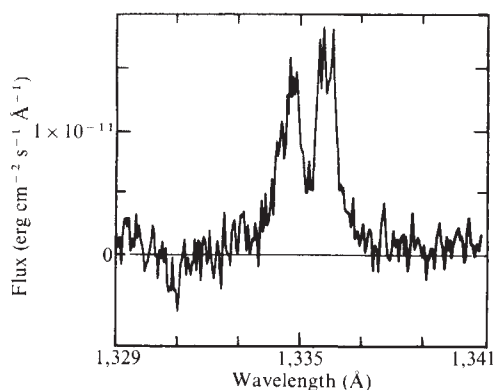


Fig. 6 Low-dispersion short-wavelength spectra of a, HR1099; b, λ And; and c, ϵ Eri in absolute flux units are compared with the quiet Sun (d) spectrum degraded to 6 Å resolution. Important lines are identified. The stellar Ly α and Si III $\lambda 1206$ lines are covered by saturated geocoronal Ly α emission.

There is, however, a significant increase in line width (FWHM) with ionisation potential—C I (0.28 Å), O I (0.47 Å), Si II (0.43 Å), C II (0.82 Å), Si III (0.69 Å), C III (0.54 Å), Si IV (0.67 Å), C IV (0.89 Å), and N V (0.71 Å). These widths, except for C I and N V, are probably accurate to $\pm 10\%$. This trend suggests that turbulent motions increase with height in the outer atmosphere as occurs in the Sun¹³. The C II and C IV lines appear to be anomalously broad compared to the trend, perhaps due to significant optical depths in these lines.

In Capella, the emission lines are all single, as would be expected at phase 77.711 d (based on the orbital elements of Batten and Erceg)³⁵ when the radial velocity difference between the two stars is very small. Also, no lines seem to show a narrow component (from Capella A) superimposed on a broader component (from the more rapidly rotating Capella B)^{36,37}. Thus we cannot ascribe the observed spectrum to a specific star on the basis of our data alone, but instead rely on previous work^{3,4} for evidence that Capella A is the emitting star.

Finally, line flux ratios for O I $\lambda 1306/\lambda 1304/\lambda 1302$ and C II $\lambda 1335/\lambda 1334$ in Capella are close to unity and similar to the solar ratios, implying large optical depths in the lines. On the other hand, the C IV $\lambda 1550/\lambda 1548$ ratio is 2.1, similar to the solar ratio, 1.8, and to the ratio of *gf* values for the lines, 2.0. We conclude that these lines must be effectively thin (not thermalised), but that the anomalous width of the lines may be due to optical-depth effects. The Si II $\lambda 1817/\lambda 1808$ ratio is 3.1 in Capella, 2.8 in ϵ Eri, and 2.4 in HR1099, compared to the *gf* value ratio, 2.0. The fact that these ratios are in excess of the optically thin value is at first sight surprising, but may be explained if the lines at 1,808.4 and 1,816.9 become optically thick whilst the 1,817.4 component remains optically thin. Photons may then be transferred from 1,804.4 to 1,817.4, thus increasing the ratio of 1,817.4 plus 1,816.9 to 1,808.4 Å.

We thank G. J. Rottman for permission to cite his solar data, and D. S. Hall for comments on the RS CVn-stars and permission to cite unpublished photometry.

1. Wright, K. O. *Publ. Dom. Astrophys. Obs.* **10**, 1 (1954).
2. Kelch, W. L. *et al. Astrophys. J.* **220**, 962 (1978).
3. Kondo, Y., Morgan, T. H. & Modisette, J. L. *Astrophys. J.* **207**, 167 (1976).
4. Dupree, A. K. *Astrophys. J. Lett.* **200**, L27 (1975).
5. Vitz, R. C., Weiser, H., Moos, H. W., Weinstein, A. & Warden, E. *Astrophys. J. Lett.* **205**, L35 (1976).
6. Dupree, A. K. in *Physique des Mouvements dans les Atmosphères Stellaires* (eds Cayrel, R. & Steinberg, M.) (Colloques Internationaux du CNRS No. 250, 1976).
7. Haisch, B. M. & Linsky, J. L. *Astrophys. J. Lett.* **205**, L39 (1976).
8. Catura, R. C., Acton, L. W. & Johnson, H. M. *Astrophys. J. Lett.* **196**, L47 (1975).
9. Mewe, R., Heise, J., Gronenschild, E. H. B. M., Brinkman, A. C., Schrijver, J. & den Boggende, A. J. F. *Astrophys. J. Lett.* **202**, L67 (1975).
10. Zirin, H. *Astrophys. J.* **208**, 414 (1976).
11. Cash, W., Bowyer, S., Garmire, G. & Riegler, G. preprint (1978).
12. Burton, W. M. & Ridgeley, A. *Solar Phys.* **14**, 3 (1970).
13. Doschek, G. A., Feldman, U., VanHoosier, M. E. & Bartoe, J.-D. F. *Astrophys. J. Suppl.* **31**, 417 (1976).

14. Code, A. D. & Meade, M. R. *Wis. Astrophys. No.* 30 (1976).
15. Kelch, W. L. *Astrophys. J.* (in the press).
16. Boggess, A. *et al. Nature* **275**, 361–369 (1978).
17. Barnes, T. G. & Evans, D. S. *Mon. Not. R. astr. Soc.* **174**, 489 (1976).
18. Eilek, J. A. & Walker, G. A. H. *Publ. astr. Soc. Pac.* **88**, 137 (1976).
19. Young, A. & Koniges, A. *Astrophys. J.* **211**, 836 (1977).
20. Walker, E. C. *J. R. astr. Soc. Can.* **38**, 249 (1944).
21. Bopp, B. W. & Fekel, F. Jr. *Astr. J.* **81**, 771 (1976).
22. Weiler, E. J. *et al. Astrophys. J.* (in the press).
23. Weiler, E. J. *Astrophys. J.* (in the press).
24. White, N. E., Sandford, P. W. & Weiler, E. J. *Nature* **274**, 569 (1978).
25. *IAU Circ. Nos* 3176, 3180, 3184, 3197.
26. Wilson, O. C. *Astrophys. J.* **138**, 832 (1963).
27. Landis, H. J. *et al. Astr. J.* **83**, 176 (1978).
28. Sandlin, G. D., Brueckner, G. E. & Tousey, R. *Astrophys. J.* **214**, 898 (1977).
29. Doschek, G. A. *et al. Astrophys. J. Lett.* **196**, L83 (1975).
30. Dupree, A. K. *Astrophys. J.* **178**, 527 (1972).
31. Haisch, B. M., Linsky, J. L., Weinstein, A. & Shine, R. A. *Astrophys. J.* **214**, 785 (1977).
32. Skelton, D. L. & Shine, R. A. *Bull. Am. astr. Soc.* **9**, 325 (1977).
33. Zirin, H. *Astrophys. J. Lett.* **199**, L63 (1975).
34. Cheng, C.-C. *Solar Phys.* **56**, 205 (1978).
35. Batten, A. H. & Erceg, V. *Mon. Not. R. astr. Soc.* **171**, 47P (1975).
36. Herbig, G. H. & Spalding, J. F. *Astrophys. J.* **121**, 118 (1955).
37. Franklin, K. L. *Astrophys. J.* **130**, 139 (1959).

IUE observations of the interstellar medium

M. Grewing

Astronomisches Institut, Tübingen, FRG

A. Boksenberg, M. J. Seaton, M. A. J. Snijders & R. Wilson

University College London, Gower St, London WC1, UK

A. Boggess, R. C. Bohlin, P. M. Perry & I. H. Schiffer III

Goddard Space Flight Center, Greenbelt, Maryland 20771

P. M. Gondhalekar

SRC Appleton Laboratory, Slough, Berks, UK

F. Macchetto

Astronomy Division, ESTEC, Noordwijk, Holland

B. D. Savage

University of Wisconsin, Madison, Wisconsin 54220

E. B. Jenkins

Princeton University Observatory, Princeton, New Jersey 08540

H. M. Johnson

Lockheed Palo Alto Research Laboratory, Palo Alto, California 94304

M. Perinotto

Osservatorio Astrofisico, Arcetri, Italy

D. C. B. Whittet

St Andrews University, St Andrews, UK

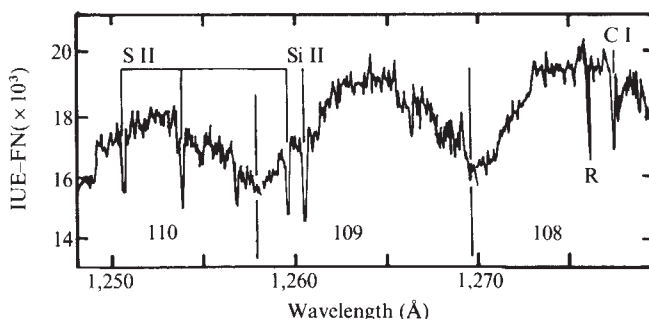
Results from a preliminary study of interstellar absorption lines in the spectra of the three stars HD149757, HD93521 and HD153919, observed by the IUE during its commissioning phase are presented. They demonstrate the capabilities of the IUE for interstellar research. Results from a preliminary study of the emission line spectrum of the planetary nebula NGC7027 are outlined and used to study its interstellar reddening.

THE IUE with its wide spectral coverage, roughly 1,150–3,200 Å embracing many of the resonance lines of the common elements, its excellent sensitivity throughout this band, and its large resolving power of 10^4 when used in the high dispersion mode, is a unique instrument for studying the interstellar medium. Atomic species that can be observed include H I, D I, Be I–II, B I–III, C I, C II, C IV, N I, N V, O I, Na I, Mg I–II, Al I–III, Si I–IV, P I–III, S I–III, Cl I, Ca I–II, Sc III, Ti I–III, V I–III, Cr I–II, Mn I–II, Fe I–II, Co I–II, Ni I–II, Cu I–II and Zn I–II. In some cases, such as C I, C II, O I, Si I, Si II, Fe II, not only the resonance lines, but also lines originating from excited fine-structure levels within the ground state are accessible. In addition to the atomic species, several diatomic and polyatomic molecules which are thought to be abundant in interstellar clouds have lines within the IUE band. Obvious candidates for such observations are OH, CO, C₂ and CH; the first, with its

transition near 3,080 Å, and the last, with its transitions near 3,137–3,146 Å, are also accessible from the ground. Some of the rarer molecular species which still might be detected in favourable circumstances include H₂O, MgH⁺, CH₃, CO⁺, NO, NO⁺, SiO and CS.

In the low dispersion mode IUE can also provide important new information about the absorption and scattering characteristics of interstellar dust.

Fig. 1 Observed spectrum of HD149757 without background and echelle ripple correction. The section of the spectrum contained in orders number 110–118 of the short-wavelength spectrograph is plotted showing strong narrow absorption lines due to interstellar C I, Si II and S II. The features indicated by R are due to fiducial marks. Vertical lines mark the ends of echelle orders.



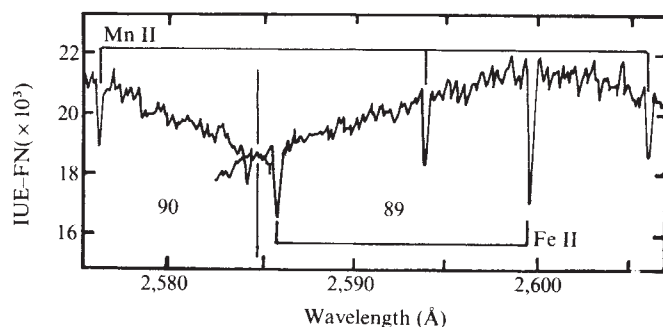


Fig. 2 Spectrum of HD149757 contained in orders number 90–89 of the long-wavelength spectrograph, showing strong absorption lines due to interstellar Mn II and Fe II.

During the commissioning phase of IUE the spectra of three stars, HD93521, HD149757 and HD153919, were obtained in order to study their interstellar absorption lines. As a fourth target the planetary nebula NGC7027 was observed in order to study its emission line spectrum. Table 1 lists these targets and the observations obtained.

The star HD149757 (ζ Oph) has been selected as a prime target for the study of interstellar absorption lines for many reasons. It is a bright source, it has been exceedingly well studied in the UV, in particular by the Copernicus satellite¹, thus providing an excellent basis to demonstrate the IUE capabilities in comparison with these other instruments. It also has a large H_2 column density, suggesting that other molecules might be found along this line of sight. As the Copernicus observations did not cover all UV wavelengths equally well, the IUE observations will extend the earlier data.

HD93521 was observed as a photometric standard as well as for the study of interstellar absorption lines. It is one of the few

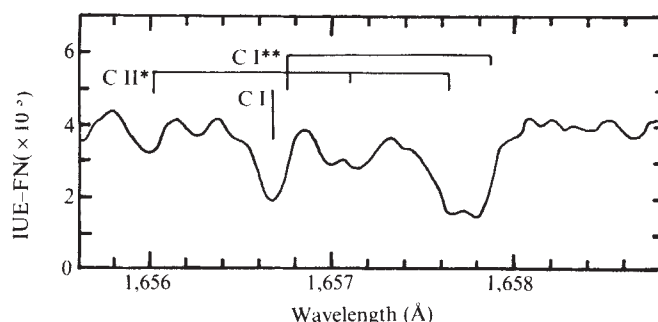


Fig. 3 The spectrum of HD149757 in the neighbourhood of the C I λ 1656.928 line, corrected for the inter-order background. Note the strength of the absorption lines originating from excited levels within the ground state of C I (* and ** indicate 1st and 2nd excited levels). The wavelength scale plotted is the one assigned during the data reduction process. The laboratory wavelengths indicated on the figure have been shifted by a constant amount of $-0.25 \text{ \AA}$ to correct for an offset in this scale.

early type stars known to be at a large distance from the galactic plane at $z = 1,570 \text{ pc}$. Also, its interstellar line spectrum shows a multiple component structure with a velocity separation of several tens of km s^{-1} (ref. 2), which is resolvable by the IUE.

HD153919 was primarily observed as it is the optical counterpart of an X-ray source (3U1700–37); and is discussed by Dupree *et al.*³ in the following article. However, it is one of the most highly reddened stars observed by IUE during the commissioning phase making it an excellent candidate for a search for interstellar absorption lines from rare species.

Absorption lines in the spectra

We now present illustrative results from a study of interstellar absorption lines in the spectra of the three stars HD149757, HD93521 and HD153919. The analysis is preliminary and based on data that had not been fully processed at the time.

HD149757

Figures 1 and 2 show two $30 \text{ \AA}$ bands at $1,250\text{--}1,280 \text{ \AA}$ and $2,575\text{--}2,605 \text{ \AA}$ extracted from the short-wavelength and long-wavelength spectrograms. The first of these covers the echelle orders 110–108 in the short-wavelength spectrograph and the data are uncorrected for the echelle ripple. The second covers the orders 90–89 in the long-wavelength spectrograph. In both cases, the strong interstellar lines from S II, Si II, Mn II and Fe II are clearly visible and have an equivalent width of the order of $100 \text{ m\AA}$. In Fig. 3 is plotted a $3 \text{ \AA}$ segment of the short-wavelength spectrograph around $1,657 \text{ \AA}$ where several C I lines occur from either the ground level proper or from excited levels within the ground state term. The equivalent width of the C I λ 1656.928 line is $\sim 70 \text{ m\AA}$. Note that the wavelength scale plotted is the one assigned during the data reduction process. It has an offset of $\Delta\lambda = 0.25 \text{ \AA}$ which has been taken into account when locating the expected line positions.

In Fig. 4 the data for five spectral lines observed by the Copernicus satellite are compared with those observed by the IUE. These lines and their equivalent widths as measured by Morton¹ are S III λ 1190.206 ($67 \text{ m\AA}$), Si II λ 1190.418 ($132 \text{ m\AA}$), P II λ 1301.87 ($15.6 \text{ m\AA}$), O I λ 1302.169 ($201 \text{ m\AA}$) and Mg II λ 2802.704 ($290 \text{ m\AA}$). The Copernicus profiles have been obtained by averaging two, three and two scans, respectively, whereas the IUE data are from a single image. Figure 4 demonstrates that the resolution of the IUE as compared to the Princeton spectrophotometer is somewhat degraded but that features separated by $0.2 \text{ \AA}$ can still be clearly resolved. Also note the much lower noise in the Mg II λ 2802 line, as observed by the IUE.

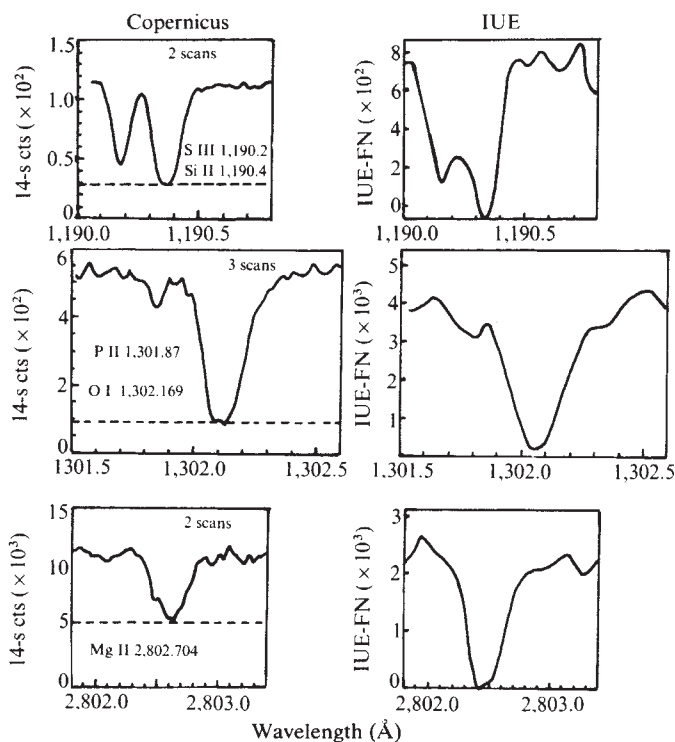


Fig. 4 Comparison of spectral line profiles as observed by the Copernicus satellite (Morton 1974) and by IUE in the spectrum of HD149757. The IUE wavelength scale has not been corrected for an offset.

Table 2 shows a comparison of the equivalent widths measured by the two systems. A statistical assessment shows that the IUE values are on average 10% higher than Copernicus and the r.m.s. deviation of the ratio of the two is $\pm 16\%$. This agreement is very satisfactory, particularly as the IUE data reduction and calibration were in their early stages and most of the large differences do occur for lines which are at present difficult to measure on IUE spectra.

Table 1 Targets for interstellar line observations

Object	$\langle V \rangle$	$\langle \text{Sp.T.} \rangle$	$\langle E(B-V) \rangle$	Image 1				IUE observations Image 2				Image 3			
				1	2	3	4	1	2	3	4	1	2	3	4
HD93521	6.89	0.9 V	0.03	SWR	HR	SA	30								
HD149757	2.56	0.9 V	0.32	SWR	HR	SA	22	LWP	HR	SA	2	LWP	HR	SA	7
HD153919	6.55	0.6 f	0.58	SWR	HR	LA	100	LWP	HR	LA	30				
NGC7027	10	—	—	SWR	LR	LA	60	LWR	LR	LA	120	LWR	HR	LA	180

* 1, Camera; 2, Spectrograph; 3, Aperture; 4, Exposure time (min).

LWP, long-wavelength prime; LWR, long-wavelength redundant; SWP, short-wavelength prime; SWR, short-wavelength redundant.

HR, high resolution; LR, low resolution. SA, small aperture; LA, large aperture.

HD93521

To study interstellar matter well above the galactic plane, the halo star HD93521 has been observed. It is located at $b^{\text{II}} = 62^\circ$ at a distance of 1,778 pc, giving $z = 1,570$ pc. The absorption lines originating from the local gas and the high z gas have a large velocity separation. The strongest components occur at $v(\text{heliocentric}) = -12 \text{ km s}^{-1}$ and -56 km s^{-1} with additional components at $+5 \text{ km s}^{-1}$ and -34 km s^{-1} (see ref. 2). For the strongest two components the H I column density is $2 \times 10^{19} \text{ cm}^{-2}$ and $8 \times 10^{19} \text{ cm}^{-2}$ (ref. 4). H₂ has been detected in the -12 km s^{-1} component⁵.

Figures 5 and 6 show the absorption line profiles as observed by IUE. For comparison the Ca II K-line profile measured by Münch and Zirin² is included in Fig. 5. In plotting the IUE results we have kept the wavelength scale as assigned during the data reduction process but shifted the marks that indicate the expected line position by -56 km s^{-1} and -12 km s^{-1} with respect to this scale. The main components are clearly seen in the C I $\lambda 1277.245$, C II $\lambda 1334.532$, C II $\lambda 1335.708$ and Si II $\lambda 1259.520$ lines. In these cases the -56 km s^{-1} component seems to be weaker than the -12 km s^{-1} one, consistent with the optical and H I results. In the case of the Si II $\lambda 1190.418$ and S III $\lambda 1190.206$ lines the same may hold but a deconvolution of

Fig. 5 Line profiles in the spectrum of HD93521. The top left plot is a reproduction of the Ca II K-line shape as observed by Münch and Zirin². The main absorption components occur at -12 km s^{-1} and -56 km s^{-1} . The expected positions of these components in the UV lines observed by IUE are shown.

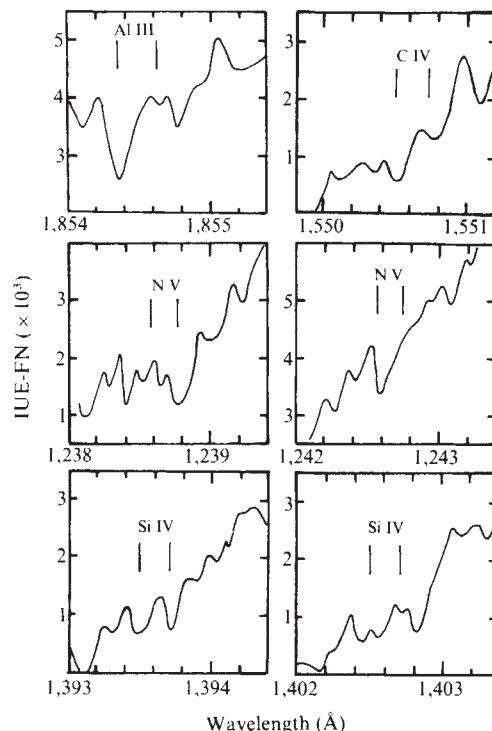
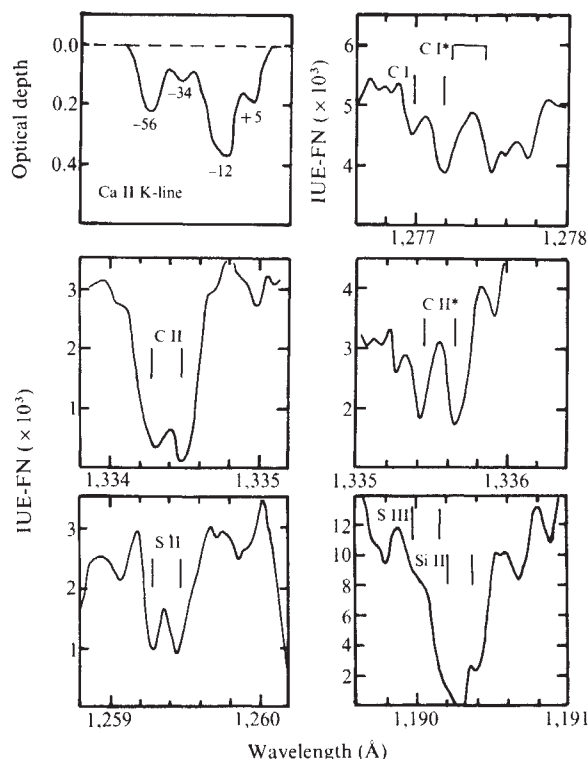


Fig. 6 As Fig. 5, except that interstellar absorption lines originating from highly ionised species have been plotted.

the observed profiles would be required to demonstrate this. An enhanced strength of the -56 km s^{-1} component which is associated with the gas at $z \geq 500$ pc (refs 2, 4 and 6) is apparent in the Al III $\lambda 1854.720$ line. The same may be true for the C IV $\lambda 1550.774$ and Si IV $\lambda\lambda 1393.755, 1402.769$ lines. However, the N V $\lambda\lambda 1238.821, 1242.804$ lines do not seem to follow this trend. Indeed there is no strong evidence for the presence of interstellar N V.

HD153919

HD153919 was observed primarily as an X-ray source and a more extensive discussion of its interstellar characteristics is given by Dupree *et al.*³ It shows the richest interstellar absorption line spectrum of the three stars observed. Not only are the lines considerably stronger than those in the spectrum of HD149757, which is expected from the higher reddening, but the lines also indicate a wider range of ionisation along this line of sight. Most noticeable is the strength of the lines arising from such highly ionised species as C IV and Si IV.

Figures 7 and 8 show two 30 Å sections of the spectrum of HD153919. Strong narrow absorption lines from singly ionised species like Si II, P II, Mn II and Fe II are clearly visible as well as the C IV lines at $1,550 \text{ Å}$. As indicated in Fig. 7, lines arising from excited levels within the ground state are strong in this spectrum. This is even more evident from Fig. 9, where we have plotted the spectrum in the neighbourhood of the C I $\lambda 1656.928$ line. Figure 10 shows the evidence for the more highly ionised

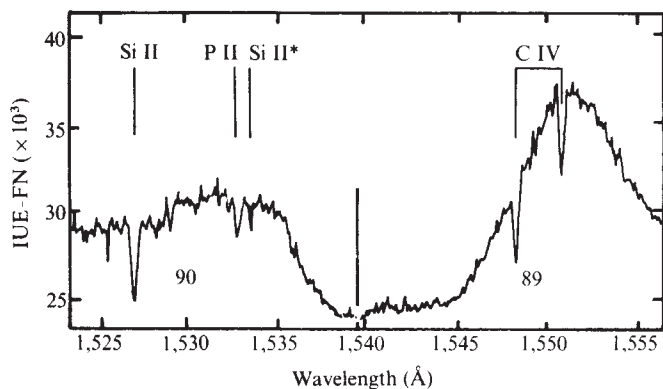


Fig. 7 Observed spectrum of HD153919 with neither background nor echelle ripple correction. The section plotted is contained in the echelle orders number 90–89 in the short-wavelength spectrograph and shows evidence for strong interstellar absorption lines due to single as well as multiple ionised species.

species such as Si III, Al III, C IV and Si IV. The presence of N V remains doubtful. X-ray ionisation of the interstellar medium local to this source might explain the presence of these high ion stages^{3,7}.

Emission lines in the spectrum of NGC7027

NGC7027 has been selected as a prime target for the study of its emission lines in the UV on several grounds. First, it was known to exhibit a strong UV line spectrum from the work of Bohlin *et al.*⁸. These data have been obtained with a rocket-borne 33 cm telescope feeding a spectrograph that yielded 46 Å mm^{-1} dispersion in the 1,300–2,900 Å region. Second, with its angular size of 10 arc s, the nebular image fits nicely in the 10×20 arc s entrance slot of the IUE spectrograph. Third, and perhaps most

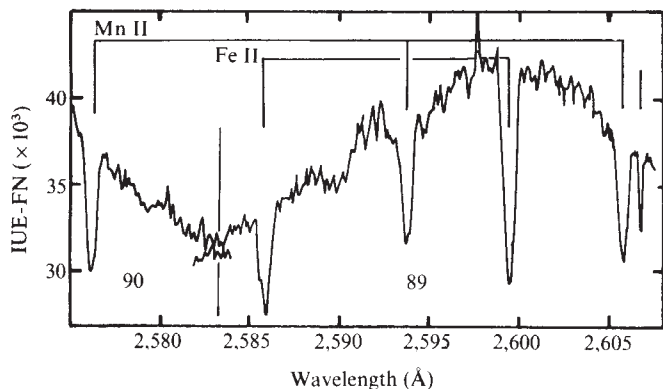


Fig. 8 As Fig. 7, except that a 30 Å section from the long-wavelength spectrograph of HD153919 has been plotted, showing the great strength of the Mn II and Fe II lines (see Fig. 2).

important, NGC7027 is now the one planetary nebula for which the emission spectrum has been studied all the way from the radio band⁹, through the near IR (10–12.5 μm (ref. 10); 4–8 μm (ref. 11); 2–4 μm (ref. 12); 0.8–2.7 μm (ref. 13)), the optical (Kaler *et al.*¹⁴ discuss the entire band from 3,050 to 8,600 Å), into the UV. These extensive data should provide an excellent basis for the detailed comparison with model calculations for this planetary nebula.

NGC7027 was observed three times with IUE. Low resolution spectra of the long-wavelength and the short-wavelength range and a high resolution spectrum of the long-wavelength range were obtained. Figure 11 shows the low resolution spectra. The intensities have not been corrected for the IUE sensitivity response but the zero point of the wavelength scale was corrected on the basis of identified lines.

The most notable feature in the spectrum is the great strength of the emission lines, most of which are readily identified. These are listed in Table 3, together with preliminary flux values. The fluxes were determined from sensitivity curves derived from observations of standard stars as described by Boggess *et al.*²³ but with the TD-1 (S2/68) fluxes as standard²⁴. The relative line intensities within each spectrum should be accurate to better than $\pm 30\%$ for the stronger features, and roughly $\pm 50\%$ for the weaker ones, except for $\lambda > 3,100 \text{ Å}$ where the sensitivity is low and the $\pm 50\%$ error applies to all lines. To combine the data from the two spectral ranges, the long-wavelength data were fitted to the short-wavelength data by using the intensity of the

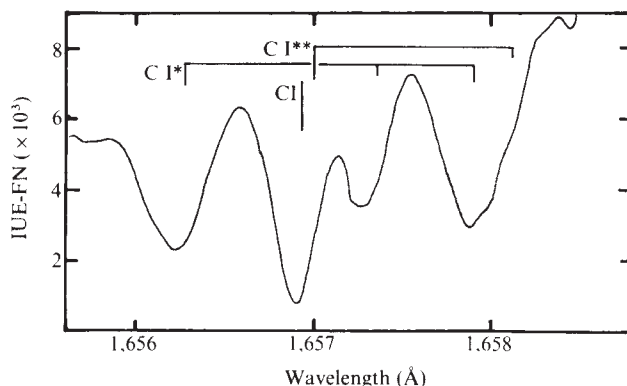


Fig. 9 The spectrum of HD153919 in the neighbourhood of the C I $\lambda 1656.928$ line. Note the great strength of the absorption lines originating from excited fine structure levels within the ground state of C I. The wavelength scale is the one assigned during the data reduction process.

C III $\lambda 1909$ line which is seen in both spectra. A comparison of these results with those from other experiments shows excellent agreement with the intensities as observed by the ANS satellite¹⁵ for the C IV $\lambda 1550$ line, and consistent agreement with the results from the rocket experiment of Bohlin *et al.*⁸, which are, however, systematically smaller, by a factor of ~ 2 . For $\lambda >$

Table 2 Equivalent widths for strong lines in the spectrum of HD149757 as measured by Copernicus (COP) and IUE

Ion	λ_{Lab} (Å)	W_{λ} (COP) (mÅ)	W_{λ} (IUE) (mÅ)	Ion	λ_{Lab} (Å)	W_{λ} (COP) (mÅ)	W_{λ} (IUE) (mÅ)
C II	1,334.532	189	224	Si II	1,193.289	147	75†
C II*	1,335.703	140	184†	Si II	1,190.416	132	130
N I	1,200.711	133	194†	Si II	1,259.520	112	148†
N I	1,200.224	131	151	Si II	1,253.812	106	112
N I	1,199.549	146	167	Si II	1,250.586	100	105
O I	1,302.169	201	198	Mn II	2,605.697	111	115
Mg I	2,852.127	218	196	Mn II	2,593.731	116	163
Mg II	2,802.704	290	337†	Mn II	2,576.107	138	137
Mg II	2,795.528	312	309	Fe II	2,599.395	226	247
Si II	1,526.708	190	156	Fe II	2,585.876	214	174
Si II	1,304.372	132	127	Fe II	2,382.034	238	259
Si II	1,260.421	170	223§	Fe II	2,373.733	120	185†

As strong lines we have arbitrarily defined lines that have an equivalent width of more than 100 mÅ as measured by Morton in the Copernicus spectrum of HD149757. Uncertainties (denoted by †) arise in the IUE data if the line is close to the end of an order or in a badly exposed part of the image. The Si II $\lambda 1,260.421$ line is blended by the Fe II $\lambda 1,260.542$ line (denoted by §).

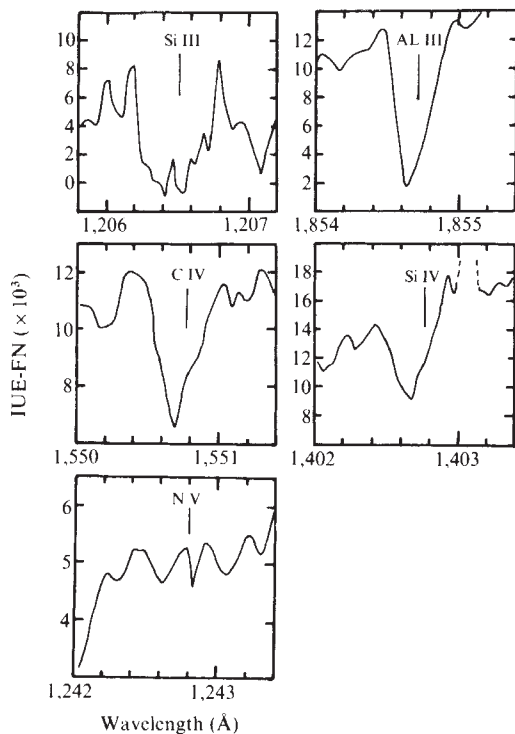


Fig. 10 Sections from the short-wavelength spectrogram of HD153919 showing evidence for interstellar absorption lines originating from highly ionised species. The raw wavelength scale is plotted.

3,100 Å, where the calibration is still rather uncertain, the ANS results and the measurements of Kaler *et al.*¹⁴ suggest that the IUE fluxes are systematically too low by 25–45%.

Figure 12a plots, on an enlarged scale, the 2,760–2,820 Å section of the low resolution spectrum showing a broad extended hump which contains the Mg II λ 2795, 2803 doublet. Figure 12b shows the same portion of the spectrum as observed by the high resolution mode. The wavelength scale is, in both cases, the uncorrected scale assigned during the reduction process. The two doublet lines as well as the [Mg v] line at λ 2783 Å are clearly seen. This example shows the potential of the high resolution mode of IUE for studying the emission from planetary nebulae in great detail.

While the low resolution short-wavelength spectrum of NGC7027 resembles in many ways the spectrum of NGC7662 as observed by Bohlin *et al.*¹⁶ in that C IV λ 1550, He II λ 1640 and C III λ 1909 are the most prominent features, the low resolution long-wavelength spectrum looks by comparison much richer in emission lines. Bohlin *et al.* quote the

Table 3 UV line fluxes in the spectrum of NGC7027

(Å)	Ion	Obs. flux (10^{-12} erg cm $^{-2}$ s $^{-1}$)	(Å)	Ion	Obs. flux (10^{-12} erg cm $^{-2}$ s $^{-1}$)
1,406	O IV]	0.8*	2,783	[Mg v]	1.0†
1,487	N IV]	0.5*	2,795	Mg II	0.1†
1,549	C IV	20.4*‡	2,803	Mg II	1.6†
1,602	[Ne IV]	1.1*	2,836	O III	0.7†
	+?				
1,640	He II	5.2*	2,853	[Ar IV]	0.2†
1,663	O III]	0.9*	2,868	[Ar IV]	0.1†
1,750	N III]	0.8*	2,928	[Mg v]	0.2†
1,909	C III]	10.8*‡	3,024	O III	0.4†
2,321	[O III]	0.9†	3,047	O III	1.5†
2,422	[Ne IV]	1.3†	3,133	O III	8.3*‡
2,470	[O II]	0.4†	3,203	He II	2.2†
2,511	He II	0.2†	3,345	[Ne v]	5.5†
2,734	He II	0.9†	3,426	[Ne v]	†

* Seen in the short-wavelength spectrum

† Seen in the long-wavelength spectrum

‡ Some pixels are saturated

Table 4 UV continuum fluxes in NGC7027

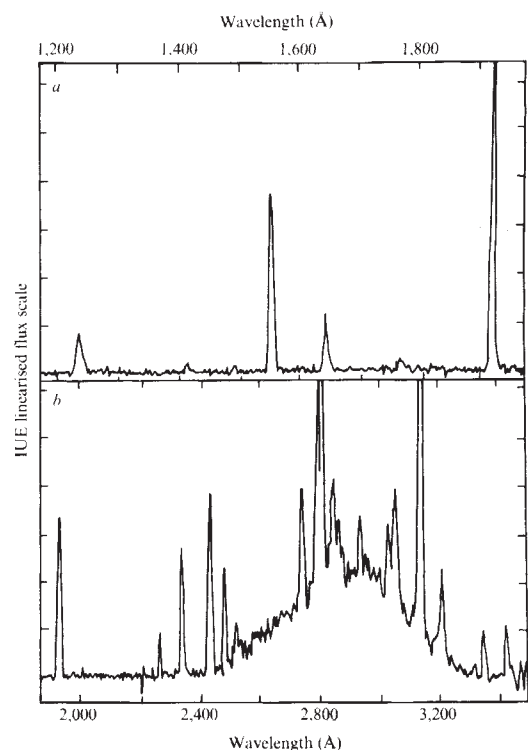
(Å)	Observed flux (10^{-15} erg cm $^{-2}$ s $^{-1}$ Å $^{-1}$)	(Å)	Observed flux (10^{-15} erg cm $^{-2}$ s $^{-1}$ Å $^{-1}$)
1,400	13	2,400	4
1,600	15	2,600	10
1,800	12	2,800	20
2,000	14	3,000	36
2,200	6	3,200	46

wavelengths of the components of the [Ne IV] $^2D_{3/2} \rightarrow ^4S$, $^2D_{5/2} \rightarrow ^4S$ doublet to be 2,439 and 2,441 Å. However, Edlén (ref. 17 and personal communication) gives these wavelengths as 2,422.60 and 2,425.24 Å, from isoelectronic sequence extrapolations. The high resolution spectrum shows the doublet to be clearly resolved, and the wavelengths to be in agreement with Edlén's prediction, to within the accuracy of the present measurements (~ 0.5 Å). The intensity ratio $R = I(\lambda 2425)/I(\lambda 2423)$ is sensitive to electron density, being similar to the well known [O II] ratio $I(\lambda 3729)/I(\lambda 3726)$. In the limit of low densities, $R = 1.5$, and in the limit of high densities, $R = 0.16$ (the latter value being obtained using the transition probabilities of Garstang¹⁸). For NGC7027 we estimate R to be about unity (see Fig. 13) which gives a value of electron density of about 5×10^4 cm $^{-3}$.

Table 4 lists preliminary estimates of the UV continuum fluxes from NGC7027. Although the modification of the intensity distribution by extinction is substantial, the data resemble the Edistribution expected from hydrogen recombination and two-photon emission.

The UV emission lines of NGC7027 can also be used in a study of interstellar absorption. This is usually investigated by comparing reddened and unreddened stars which are believed to be intrinsically identical from their line spectra. Planetary nebulae provide a totally different method in that the recombination lines in H I and He II can be linked directly to the free-free radio emission. Because the interstellar medium is optically thin to the latter, its measured flux gives the intrinsic

Fig. 11 The low resolution *a*, short- and *b*, long-wavelength spectra of NGC7027 as observed by IUE. The flux values have not been corrected for the instrumental sensitivity. For an identification of the emission lines and corrected flux values see Tables 3 and 4.



fluxes in the H I and He II lines. A comparison with the observed line fluxes then gives directly the interstellar extinction.

NGC7027 is particularly appropriate to such a study because it is heavily reddened and has been observed extensively. Previous studies by Miller and Matthews¹⁹ used radio data and the H α , H β and H γ lines, and by Schwartz and Peimbert²⁰ who included lines at $\lambda = 1 \mu\text{m}$. With the observation of the UV He II lines with IUE an investigation of interstellar extinction over a very wide wavelength range is now possible and a brief analysis is given here. A more extensive and detailed discussion will be given elsewhere²¹.

The values of the interstellar extinction in NGC7027— $A_\lambda(7027)$ —have been derived from the IUE observations of He II lines on the basis of the method outlined above and combined with those obtained from ground-based data. The results are plotted in Fig. 14 against the values of A_λ/E_{B-V} derived by Nandy *et al.*²² from a study of early-type stars with the sky survey telescope (S2/68) on the TD-1 satellite. The excellent straight line fit of the data shows that over a wide range of wavelengths, from H α 6,563 Å to He II 1,640 Å, the reddening of NGC7027 follows closely the standard reddening law as derived by Nandy *et al.*²².

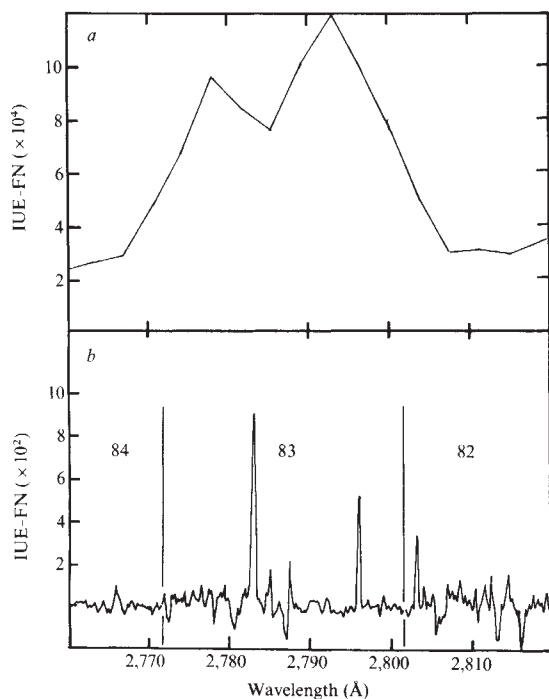


Fig. 12 Comparison of *a*, low resolution and *b*, high resolution spectra of NGC7027 in the neighbourhood of the Mg II λ 2800 doublet. The high resolution spectrum is extracted from the echelle orders number 84–82. The wavelength scale is the one assigned during the data reduction process.

For the longer wavelengths there is some discrepancy between the extinction deduced for NGC7027 and the standard extinction. The straight line of Fig. 14 does not pass through the origin and can be represented by

$$A_\lambda(7027) = \alpha \left(\frac{A_\lambda}{E_{B-V}} \right) + \beta$$

where $\alpha = 0.935$ and $\beta = 0.17$. The difference is small but there is no doubt that it is real.

An explanation of this discrepancy can be made by considering the large variation of extinction across the face of the nebula which is known to exist from the quite different distributions in the radio and optical images. Let us consider a simple two-component model in which there is a heavy local extinction over a fraction B of the total area of the source. We take the

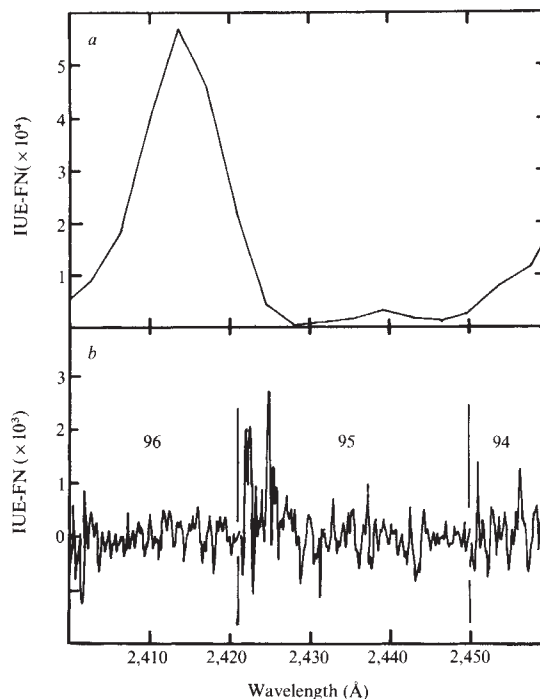


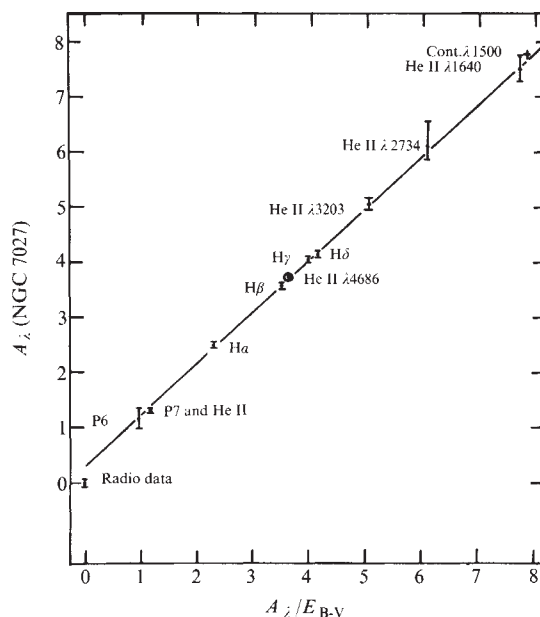
Fig. 13 As Fig. 12, except that the spectrum in the neighbourhood of the [Ne IV] doublet has been plotted.

extinction from area B to be $\gamma(A_\lambda/E_{B-V})$ and from the remaining $(1-B)$ to be $\alpha(A_\lambda/E_{B-V})$ with $\alpha = 0.94$ as derived above and $\gamma \gg \alpha$. We can then write

$$10^{-0.4 A_\lambda(7027)} = B 10^{-0.4 \gamma (A_\lambda/E_{B-V})} + (1-B) 10^{-0.4 \alpha (A_\lambda/E_{B-V})}$$

Assuming the first term to be negligible over the wavelength range 6,563–1,640 Å where the straight line fit is good, then the two-model equation is consistent with that fit if $(1-B) = 10^{-0.4 \beta}$, giving $B = 0.145$. The value of γ can now be deduced by fitting in the red and IR regions. Using the data plotted in Fig. 14 together with the most extreme IR measurement yet available (Brackett γ at 2.166 μm , Merrill *et al.*¹²), a good fit is obtained with $\gamma = 2.0$. Hence, this simple, two component model gives

Fig. 14 Interstellar extinction measured for NGC7027 in the radio, IR, optical and UV ranges. A_λ for NGC7027 has been plotted against $\{A_\lambda/E_{B-V}\}$ as determined by Nandy *et al.*²² from observations of about 100 reddened early type stars with the Sky Survey Telescope (S2/68) in the ESO TD 1 satellite.



excellent agreement with the observations over the full wavelength range from the IR to the UV.

The reddening derived for NGC7027 over the range 6,563–1,640 Å is in excellent agreement with that derived from early type stars by Nandy *et al.* Most of the reddening is probably due to interstellar grains along the line of sight. The results for larger wavelengths can be explained by a local extinction, variable over the nebula, caused by grains with the same optical properties.

1. Morton, D. C. *Astrophys. J.* **197**, 85 (1975).
2. Münch, G. & Zirin, H. *Astrophys. J.* **133**, 11 (1961).
3. Dupree, A. K., *et al.* *Nature* **275**, 384–387 (1978).
4. Habing, H. *Bull. astr. Inst. Netherlands* **20**, 128 (1969).
5. Savage, B. D., Bohlin, R. C., Drake, J. F. & Budich, W. *Astrophys. J.* **216**, 291 (1977).
6. Rickard, J. J. *Astr. Astrophys.* **17**, 425 (1972).
7. McCray, R., Wright, Chr. & Hatchett, S. *Astrophys. J. Lett.* **211**, L29 (1977).

8. Bohlin, R. C., Maronni, P. A. & Stecher, Th. P. *Astrophys. J.* **202**, 415 (1975).
9. Churchwell, E., Terzian, Y. & Walmsley, M. *Astr. Astrophys.* **48**, 331 (1976).
10. Aitken, D. K. & Jones, B. *Mon. Not. R. astr. Soc.* **165**, 363 (1973).
11. Russell, R. W., Soifer, B. T. & Willner, S. P. *Astrophys. J. Lett.* **217**, L149 (1977).
12. Merrill, K. M., Soifer, B. T. & Russell, R. W. *Astrophys. J. Lett.* **200**, L37 (1975).
13. Treffers, R. R., Fink, U., Larson, H. P. & Gautier III, T. N. *Astrophys. J.* **209**, 793 (1976).
14. Kaler, J. B., Aller, L. H., Czyzak, S. J. & Epps, H. W. *Astrophys. J. Suppl. Ser.* **31**, 163 (1976).
15. Pottasch, S. R., Wesselius, P. R., Wu, C.-C., Fieten, H. & van Duinen, R. J. *Astr. Astrophys.* **62**, 95 (1978).
16. Bohlin, R. C., Harrington, J. P. & Stecher, Th. P. *Astrophys. J.* **219**, 575 (1978).
17. Edlén, B. *Solar Phys.* **24**, 356 (1972).
18. Garstang, R. H. *Planetary Nebulae* (eds Osterbrock, J. E. & O'Dell, C. R.) 143 (Reidel, Dordrecht, 1968).
19. Miller, J. S. & Mathews, W. G. *Astrophys. J.* **172**, 593 (1972).
20. Schwartz, R. D. & Peimbert, M. *Astrophys. Lett.* **13**, 157 (1973).
21. Seaton, M. J. *Mon. Not. R. astr. Soc.* (submitted).
22. Nandy, K., Thompson, G. I., Jamar, C., Monfils, A. & Wilson, R. *Astr. Astrophys.* **44**, 195 (1975).
23. Boggess, A. *et al.* *Nature* **275**, 361–369 (1978).
24. Boksenberg, A. *et al.* *Nature* **275**, 388–391 (1978).

IUE observations of X-ray sources: HD153919 (4U1700–37), HDE226868 (Cyg X-1), HZ Her (Her X-1)

**A. K. Dupree, R. J. Davis, H. Gursky,
L. W. Hartmann & J. C. Raymond**

Harvard-Smithsonian Center for Astrophysics, Cambridge,
Massachusetts 02138

A. Boggess, A. Holm, Y. Kondo & C.-C. Wu

Goddard Space Flight Center, Greenbelt, Maryland 20771

F. Macchetto

Astronomy Division, ESTEC, Noordwijk, Holland

M. C. W. Sandford

SRC, Appleton Laboratory, Slough, Berks, UK

A. J. Willis & R. Wilson

University College London, Gower Street, London WC1, UK

F. Ciatti

Osservatorio Astrofisico, I-36012 Asiago, Italy

J. B. Hutchings

Dominion Astrophysical Observatory, Victoria, BC, Canada V8X 3X3

H. M. Johnson

Lockheed Palo Alto Research Laboratory, Palo Alto, California 94304

J. Jugaku

Tokyo Astronomical Observatory, University of Tokyo, Mitaka-shi,
Tokyo, Japan

D. C. Morton

Anglo-Australian Observatory, Epping N.S.W. 2121, Australia

A. Treves

Laboratorio di Fisico Cosmico del C.N.R., Milano, Italy

E. P. J. van den Heuvel

Astronomical Institute, Amsterdam, Holland

The first UV spectroscopic measurements of binary X-ray sources with IUE show highly variable emission from a photoionised plasma in the object HZ Her, and give evidence for localised circumsystem material in the binary source Cygnus X-1. In addition, a substantial stellar wind is found in one of the brightest identified X-ray sources HD153919. This system may be surrounded by an extensive H II region unusual in its content of high excitation ion species.

BINARY X-ray sources containing a compact object represent a class of stars that generally have been inaccessible to UV spectroscopic observations due to their faintness. Studies of these objects can potentially discern the physical conditions and dynamics of the material in the sources themselves as well as probe the interaction between the source and the surrounding interstellar medium. The IUE was used during the high priority phase of the commissioning period to obtain spectra of three stars—HD153919, HDE226868, HZ Her—that correspond to the X-ray sources 4U1700–37, Cyg X-1, and Her X-1 respectively. We report here some general characteristics of the spectra, comments on the presence of phase effects, and note the

strength of interstellar lines towards one source, HD153919, where high resolution spectra are available.

Observations

A discussion of the instrument and its performance is contained in preceding articles^{1,2}. The spectra discussed here were taken over a 7-week period beginning 12 February 1978 (see Table 1). All exposures were made with the large aperture. Corrections for the effects of the radiation background and scattered light were applied by subtracting a smoothed value of the inter-order fluxes from the spectrum in each order. The adequacy of this correction for the high dispersion spectra is verified by the appearance of strong interstellar lines (that is, H Ly α at λ 1216 and C II at λ 1335) and absorption components of P Cygni profiles with flat cores of approximately zero net flux. The high dispersion spectra of HD153919 have not been corrected for the grating blaze (echelle ripple); the low dispersion spectra have been calibrated by using the preliminary determination of the instrumental sensitivity with wavelength².

HD153919 = 4U1700–37

This binary system identified with the X-ray source⁷ is composed of a primary star classified from the optical spectrum as O6f, and a compact object—the X-ray source—orbiting essentially within

the expanding atmosphere of the primary star⁸. This source was observed at high dispersion at two phases ~ 0.35 and 0.97 , in its 3.4-d binary period.

The spectrum of the primary star shows the characteristic rich array of P Cygni profiles that has been found in other Of stars (see refs 9, 10). The lines of the strong resonance doublets of C IV ($\lambda 1550$) and Si IV ($\lambda 1400$) are shown in Figs 1 and 2 where the broad extent of these lines is apparent. Transitions that show P Cygni features and some characteristics of their profiles at phase 0.97 are given in Table 2. Here the velocity scale has been reduced to the centre-of-mass velocity of the system (-60 km s^{-1}) by using interstellar lines of low excitation and assigning them the measured heliocentric velocity of the Ca II interstellar lines¹¹, namely -14.7 km s^{-1} . The positions of line maxima were measured from tracings on a scale of $4 \text{ \AA}$ per inch; measurements to $\pm 0.1 \text{ \AA}$ can be achieved. A greater uncertainty arises in selecting the appropriate positions on the profiles since effects of the grating blaze have not been removed from these scans. This can cause a shift in the apparent maximum of a P Cygni profile and errors in the assignment of the shortwards edge. An uncertainty of $\pm 100 \text{ km s}^{-1}$ is probably a reasonable estimate of the errors in this procedure. The observed terminal velocities $\sim 2,600 \text{ km s}^{-1}$ are comparable to those found in ζ Pup (ref. 12) (spectral type O4 If) and are larger than velocities occurring in other O supergiants of later spectral type¹⁰, for example, 29 C Ma (O7 Ia:f/p), 9 Sge (O7.5Iaf), and HD151804 (O8 Iaf).

The escape velocity from HD153919 is $\sim 720 \text{ km s}^{-1}$ assuming⁸ $M = 27 M_{\odot}$ and $R = 20 R_{\odot}$. This value is less than the maximum expansion velocities found in the P Cygni profiles indicating the presence of mass loss similar to that found in other luminous stars¹⁰.

A search was made for variations in the line profiles with binary phase. It can be argued that the structure of the wind changes when the X-ray source is present, either through ionisation of the dominant ion species that drive the wind, or by ionisation of a cavity in the expanding atmosphere of the primary star that would preferentially be in the radial direction.

The two high dispersion short wavelength spectra were taken at phase 0.32 and phase 0.97, the latter when the compact object is eclipsed. Superimposition of the broad profiles of the resonance lines Si IV, N V, and C IV shows effectively the same shape at the two phases in contradiction to some predictions¹³.

The effect of the X-ray source, if any, on the atmosphere of the primary must be confined to a region which is small compared with the extended line-forming region of the resonance transitions. Transitions from excited levels that are formed over a narrow range in velocity may be more responsive to the effects of the X-ray source. An optimum time to obtain spectra may be at phase 0.5 when the compact object is directly in front of the primary star.

The interstellar spectrum is also discussed in the preceding article by Grewing *et al.*¹⁵. It is rich as might be expected from a

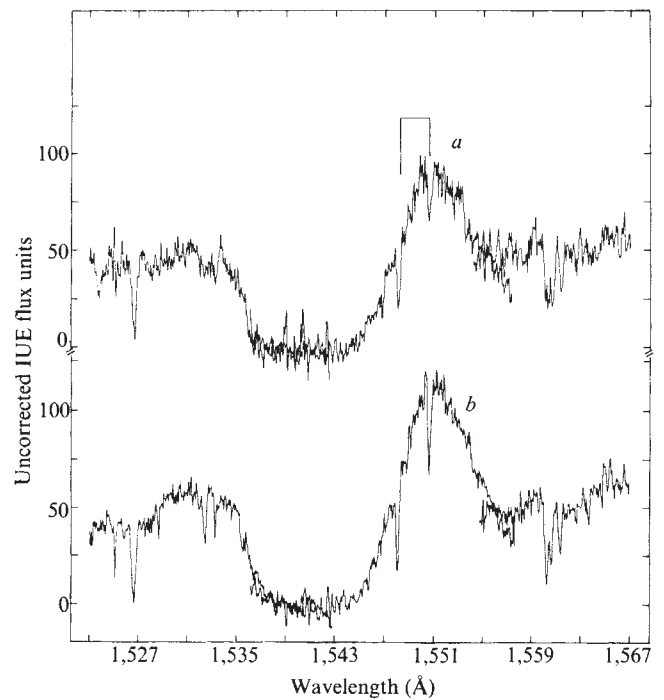


Fig. 1 The region $\lambda\lambda 1523\text{--}1567$ of HD153919 (4U1700-37) taken in the high dispersion mode at two orbital phases: *a*, phase 0.32 (SWR 1027); *b*, phase 0.97 (SWR 1054). The broad stellar P Cygni profile of the C IV doublet near $\lambda 1550$ has the same appearance and extent at both phases. The strong narrow features marked at $\lambda 1548$ and $\lambda 1550$ are attributed to interstellar C IV. Other interstellar lines arising from neutral and singly ionised species are apparent; the absorption feature at $\lambda 1526.7$ is most probably interstellar Si II. There was a change of system focus between the two images; the resolution of the lower (phase 0.97) spectrum is superior to the other, as can be noted by inspection of the interstellar profiles. A correction has not been implied for the echelle blaze; the wavelength scale is given directly as inferred from the calibration lamp. The wavelength regions where the orders overlap can be noted by the double tracings; the sharp emission features near $\lambda 1539$ are of instrumental origin.

reddened source with $E(B - V) = 0.52$ (ref. 14) at a distance of 1.2 kpc (ref. 8). Many atomic species C I, Si I, Cl I are present, reflecting a large amount of cold neutral material along the line of sight. The relative weakness of C I lines originating from excited fine-structure levels suggests that the temperature is less than 50 K (J. B. Black, personal communication).

The most notable interstellar features are the strong absorption lines of the highly ionised species of Si IV and C IV (see Figs 1 and 2). The equivalent widths for the C IV doublet at $\lambda 1550.6$ and $\lambda 1548.0$, and the Si IV transition $\lambda 1402.37$ are given in Table 3. Only one of the Si IV doublet transitions is found due to the presence of a strong stellar P Cygni profile that completely absorbs the background stellar continuum. The region near

Table 1 Summary of observations

Star	X-ray source	Camera image	Dispersion	Exposure (min)	Exposure midpoint (JD 244+)	Phase
HD153919	4U1700-37	SWR 1027	High	100	3,552.488	0.32*
		LWP 1029	High	30	3,552.624	0.36
		SWR 1054	High	100	3,564.927	0.97
HZ Her	4U1656+35 (Her X-1)	SWR 1051	Low	180	3,563.290	(1.7 day [†]) 0.82
		SWR 1056	Low	240	3,565.276	0.98
		SWP 1272	Low	90	3,598.903	0.76
HDE226868	4U1956+35 (Cyg X-1)	SWR 1050	Low	150	3,562.957	0.48‡
		SWP 1273	Low	40	3,599.049	0.93

* Ephemeris T. Matilsky (personal communication).

† Ephemeris from refs 3-5.

‡ Ephemeris from ref. 6.

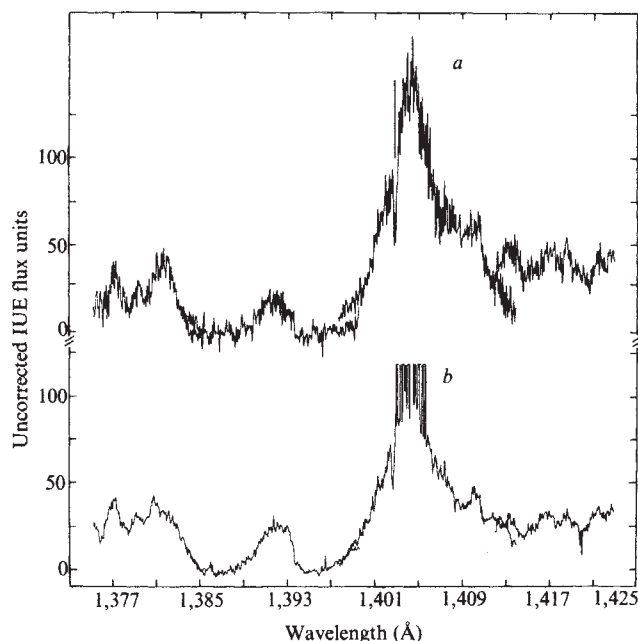


Fig. 2 The region $\lambda\lambda 1377$ – 1425 of HD153919 containing the resonance doublet of Si IV at $\lambda 1394$ and $\lambda 1403$. These images were obtained at two orbital phases of the system: *a*, phase 0.32 (SWR 1027); *b*, phase 0.97 (SWR 1054). The focus of the instrument was superior when the lower spectrum was obtained; the emission peak near $\lambda 1405$ is saturated in the lower spectrum. The stellar P Cygni profiles indicate outflow velocities of $\sim 2,500 \text{ km s}^{-1}$; there are no substantial differences in the profiles between the two spectra. Interstellar absorption due to Si IV is marked near $\lambda 1403$. See the legend to Fig. 1 for additional notes.

$\lambda 1240$ where interstellar N V occurs was not well exposed in our images, and upper limits to the equivalent widths are estimated from measurements of the peak-to-peak noise over widths of $0.4 \text{ \AA}$ at the expected position of the lines.

The lines are assumed to lie on the linear part of the curve of growth. For Si IV, in which one line is found, this assumption gives a lower limit to the column density. The C IV transitions have central intensities which are about half of the continuum intensity; moreover their equivalent widths are in the expected 2:1 ratio confirming the assumption of unsaturated lines. The derived column densities (see Table 3) are substantially higher than those found for ζ Oph, a cooler less luminous star of spectral type O9.5 V. Highly ionised species are about twice as abundant towards HD153919 than towards ζ Pup, an O4 If star of comparable spectral type. HD153919 is at a distance of 1.2 kpc which is about 2.5 times the distance of 450 pc to ζ Pup so that we cannot rule out a large scale distribution of C IV and Si IV. However, if these ion species are confined to an H II region near the star, their higher column density may reflect the

Table 2 Velocities (km s^{-1}) of P Cygni profiles in HD153919 (Phase 0.97)

Ion	λ_{lab} ($\text{\AA}$)	Emission maximum	Absorption centre	Short edge
He II	1,640.33	–3	–420	–640
C III	1,175.7	+580	–1,400	–2,200
C IV	1,550.77 1,549.06* 1,548.18	+87†	–1,500	–2,600
N IV	1,718.55	+320	–1,000	$\approx$ –1,800
N V	1,242.80 1,240.15* 1,238.82	+200	–1,400	–2,300
O IV	1,338.61	+230	–220	–440
Si IV	1,402.77 1,393.76	+270 —	–1,600 –1,500	–2,100 –2,600

* Mean wavelength for unresolved multiplet.

† This maximum may appear at shorter wavelengths due to the (uncorrected) blaze of the grating.

Table 3 Highly ionised interstellar species

Ion	λ ($\text{\AA}$)	W (m $\text{\AA}$)	f	$\log N(\text{cm}^{-2})$			
				HD153919 observed	HD153919 predicted*	ζ Pup†	ζ Oph†
C IV	1,550.77	112	0.097	13.71	14.3	13.37	<12.7
	1,548.18	213	0.194	13.72			
N V	1,238.82	<133	0.152	<13.81	13.7	<12.2	<12.3
	1,242.80	<56	0.076	<13.74			
Si IV	1,402.77	104	0.262	≥ 13.36	14.2	13.1	12.7

* Ref. 16.

† Ref. 12.

‡ Ref. 17.

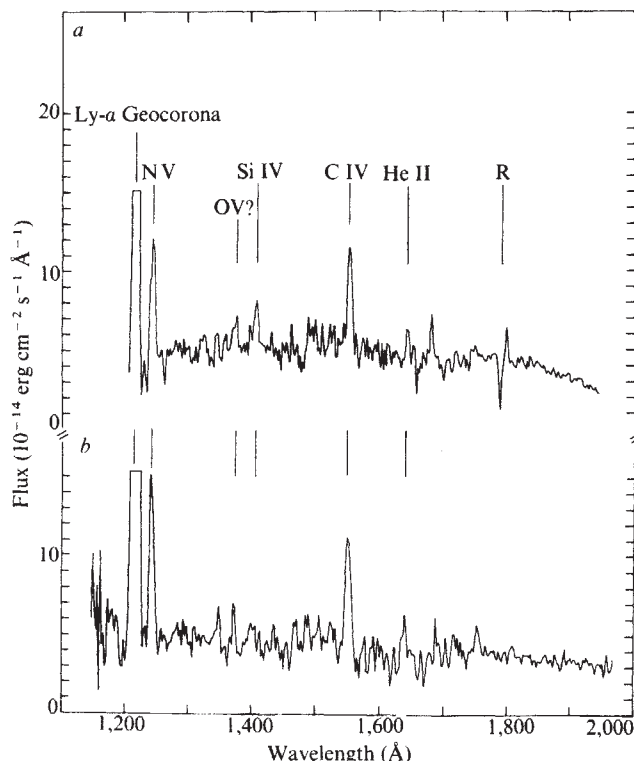
effects of ionisation by the X-ray source, although the observed column densities are still considerably less than those predicted by McCray *et al.*¹⁶ for HD153919.

Relative abundances of ion species may be used to discriminate between various sources of ionisation. In Table 4, we list the abundances of N V and C IV relative to Si IV that are observed towards HD153919 and towards ζ Oph and ζ Pup. These values are compared to calculations for a model of HD153919 immersed in the photoionised interstellar gas¹⁶, and for an interstellar bubble that can be caused by a strong stellar wind¹⁸. In the former case, a relatively cool but highly ionised gas can result from the EUV and X-ray radiation from HD153919. The bubble produces a conduction-dominated interface between the hot shocked (10^6 K) gas and the cold interstellar medium. This interface is the source of the absorption lines. The relative abundances that are observed (Table 4) are substantially less than those estimated for an interstellar bubble. The observations seem to support the presence of a photoionised gas near both the X-ray source and the two hot stars.

HZ Her (Her X-1)

The eclipsing binary system HZ Her contains a primary star that varies between spectral types A and F, and one of the first pulsating X-ray sources to be identified¹⁹. Periodic flux varia-

Fig. 3 The short wavelength spectrum of HZ Her at low dispersion ($\sim 6 \text{ \AA}$): *a*, phase 0.76; *b*, phase 0.82. Several emission lines are apparent at both phases: N V; Si IV; C IV; and He II. The feature marked R results from a fiducial mark (reseau) on the faceplate of the camera tube.



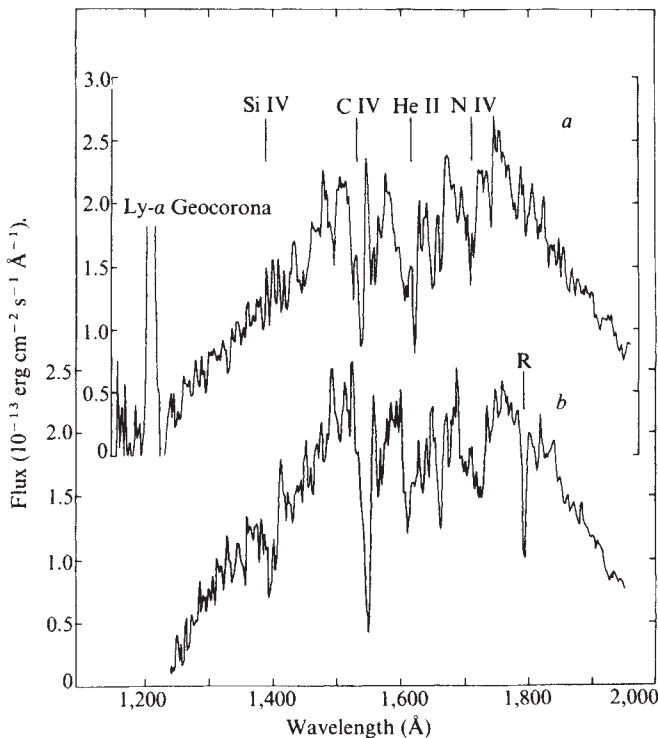


Fig. 4 The short wavelength spectrum of HDE226868 (Cyg X-1) at two phases: a, phase 0.48; b, phase 0.93. Several strong absorption lines are noted. The feature marked R near $\lambda 1800$ results from a reseau mark.

tions are present at optical and X-ray wavelengths resulting from the binary motion of the system (1.7-d period), the rotation of the compact object and perhaps X-ray shadowing of the accretion disk²⁰ (35-d X-ray on/off period, and associated optical variability). X-ray heating of the primary star, mass accretion by the compact object, and variable absorption^{21,22} cause additional fluctuations in measured emissions.

The spectrum of HZ Her from $\lambda\lambda 1200$ – 2000 was obtained at three phases during the X-ray on period of its 35-d cycle (see Table 1). Emission lines attributed to N v, Si iv, C iv and He II are clearly apparent in exposures taken at orbital phase 0.82 and 0.76 (see Fig. 3). However, a 4-h exposure centred at phase 0.98 (eclipse of the compact object) showed no line emission except for the geocoronal Ly α , suggesting that the source of UV emission is restricted in size and is located in the region on or about the opposing side of the primary, facing the compact object. An upper limit for the line and continuous flux between $\lambda 1200$ – 2000 at phase 0.98 is estimated to be 1×10^{-14} erg cm $^{-2}$ s $^{-1}$ Å $^{-1}$.

The N v emission is surprisingly strong relative to C iv. Calculations²³ have been made of the line emission from a spherical, gas cloud of uniform density that is photoionised by a central X-ray source. The calculations suggest that generally C iv should be about six times stronger than N v, while our observations suggest nearly equal emission fluxes.

There is an Fe XII line at $\lambda 1242$ which could be blended with the N v doublet at $\lambda 1240$. If so, another Fe XII line at $\lambda 1350$ should be present although weaker than $\lambda 1242$, by a factor of ~ 3 in the high density limit. A feature at $\lambda 1350$ is apparent in both spectra of HZ Her (Fig. 3); if this line is indeed Fe XII, then the contribution of Fe XII at $\lambda 1242$ to the N v blend could be significant in reducing the inferred intensity of N v.

The continuum has an approximate spectral index $F_\nu \propto \nu^{-1.3}$, and is presumably due to the heating of the primary atmosphere by X rays. This is apparently in general agreement with the calculations of Milgrom and Salpeter²⁵, although their results are presented in wide bandpasses, making direct comparison difficult. The continuum flux at $\sim 1,500$ Å is $\sim 20\%$ higher at phase 0.76 than at phase 0.82; this is a marginal result, due to the

Table 4 Relative abundances of highly ionised interstellar species

Source	N(C iv)/N(Si iv)	N(N v)/N(Si iv)
HD153919 (observed)	<2.3	<2.5
ζ Oph	<0.92	<0.38
ζ Pup	1.5–2.3	<0.08
Photoionised volume ¹⁶	1.2	0.3
Interstellar bubble ¹⁸	15.9	6.3

inherent errors and the fact that for operational reasons two different cameras were used. Milgrom and Salpeter²⁵ predict a factor of two change in the continuum between these phases, which may be ruled out.

HDE226868 (Cyg X-1)

The binary system HDE226868 is identified with the X-ray source Cyg X-1. The primary star is classified²⁶ as O9.7Iab, and the X-ray producing secondary is perhaps the strongest candidate for a black hole of any compact X-ray source.

The low dispersion spectrum of HDE226868 is shown in Fig. 4 at two phases. The spectrum from the supergiant has been detected and is generally typical of a highly reddened star. This star shows broad absorption features most probably of C iv ($\lambda 1550$), He II ($\lambda 1640$), and N iv ($\lambda 1718$). The absorption, particularly of C iv, seems to be asymmetric and indicative of mass loss. Strong P Cygni emission components as found in stars of similar type from Copernicus¹⁰ are absent. There is apparently a substantial phase effect in this spectrum with the strengthening of the Si iv and C iv transitions near superior conjunction when the supergiant lies between the Earth and the compact object. These features may be similar to the absorption dips reported in soft and hard X-ray observations that occur preferentially near superior conjunction, and reflect the presence of matter between the binary components^{28–30}. Other lines such as the feature near $\lambda 1710$ may appear at phase 0.93. We should point out that the second spectrum was taken with a different camera and a better focus, so that the depth of the features could be affected.

The strength of Si iv in early-type stars seems to be a strong function of luminosity class. The equivalent width of the Si iv doublet at phase 0.93 is 6 Å, which is less than the O9.5 Ia star α Cam, but clearly more than the O9.5 V star δ Ori A (ref. 31). Thus the strength is compatible with a high luminosity for the primary.

We thank J. H. Black for contributions to the interstellar line analysis and P. Boynton and R. McCray for useful discussions.

- Boggess, A. *et al.* *Nature* **275**, 356–361 (1978).
- Boggess, A. *et al.* *Nature* **275**, 361–369 (1978).
- Forman, W., Jones, C. A. & Liller, W. *Astrophys. J. Lett.* **177**, L103 (1972).
- Becker, R. H. *et al.* *Astrophys. J.* **214**, 879 (1977).
- Giacconi, R. *et al.* *Astrophys. J.* **184**, 227 (1977).
- Kemp, J. C. *IAU Circ.* No. 3149 (1977).
- Jones, C. *et al.* *Astrophys. J. Lett.* **181**, L43 (1973).
- Hutchings, J. B. *X-Ray Binaries* 531 (NASA SP-389, 1976).
- Morton, D. C. *Astrophys. J.* **203**, 386–398 (1976).
- Snow, T. P., Jr. & Morton, D. C. *Astrophys. J. Suppl.* **32**, 429–465 (1976).
- Hutchings, J. B. *Astrophys. J.* **192**, 677 (1974).
- Morton, D. C. *Astrophys. J.* **197**, 85–115 (1975).
- McCray, R. & Hatchett, S. *Astrophys. J.* **199**, 196 (1975).
- Hammerschlag-Hensberge, G. & Wu, C. C. *Astr. Astrophys.* **56**, 433 (1977).
- Grewing, *et al.* *Nature* **275**, 378–384 (1978).
- McCray, R., Wright, C. & Hatchett, S. *Astrophys. J. Lett.* **211**, L29 (1977).
- Morton, D. C. *Astrophys. J.* **222**, 863 (1978).
- Weaver, R., McCray, R., Castor, J., Shapiro, P. & Moore, R. *Astrophys. J.* **218**, 377 (1977).
- Tananbaum, H. *et al.* *Astrophys. J.* **174**, L143 (1972).
- Gerend, D. & Boynton, P. E. *Astrophys. J.* **209**, 562 (1976).
- Bahcall, J. N. 1975, Weizmann Institute of Science preprint no. WIS-75/35-Ph; *Proc. Enrico Fermi School on the Physics and Astrophysics of Neutron Stars and Black Holes*, Varenna, Italy, 1975 (in the press).
- Gorenstein, P. & Tucker, W. H. A. *Rev. Astr. Astrophys.* **14**, 373 (1976).
- Hatchett, S., Buff, J. & McCray, R. *Astrophys. J.* **206**, 847 (1976).
- Flower, D. R. *Astr. Astrophys.* **54**, 163 (1977).
- Milgrom, M. & Salpeter, E. E. *Astrophys. J.* **196**, 583 (1975).
- Walborn, N. R. *Astrophys. J. Lett.* **179**, L123 (1973).
- Blumenthal, G. R. & Tucker, W. H. A. *Rev. Astr. Astrophys.* **12**, 23 (1974).
- Holt, S. S., Boldt, E. A., Serlemitsos, P. J. & Kaluzienski, L. J. *Astrophys. J.* **203**, L63 (1976).
- Murdin, P. in *X-Ray Binaries* 425 (NASA SP-389, 1976).
- Parsignault, D. R., Grindlay, J., Schnopper, H., Schreier, E. J. & Gursky, H. in *X-Ray Binaries* 429 (NASA SP-389, 1976).
- Snow, T. P. & Jenkins, E. J. *Astrophys. J. Suppl.* **33**, 269 (1977).

IUE observations of extragalactic objects

A. Boksenberg, M. A. J. Snijders & R. Wilson

University College London, Gower Street, London WC1, UK

P. Benvenuti, J. Clavell, F. Macchetto & M. Penston

ESA Satellite Tracking Station, Villafranca Del Castillo, Madrid, Spain

A. Boggess & T. R. Gull

Goddard Space Flight Center, Greenbelt, Maryland 20771

P. Gondhalekar

SRC Appleton Laboratory, Slough, Berks, UK

A. L. Lane

Jet Propulsion Laboratory, Pasadena, California 91103

B. Turnrose & C. C. Wu

Computer Sciences Corporation, Silver Spring, Maryland 20900

W. M. Burton

Astrophysics Research Division of Appleton Laboratory, Culham, Oxon, UK

A. Smith

University of Florida, Gainesville, Florida 32720

F. Bertola & M. Capaccioli

Osservatorio Astronomico, Padova, Italy

A. M. Elvius

Stockholm Observatory, Sweden

R. Fosbury, M. Tarenghi & M.-H. Ulrich

European Southern Observatory, Geneva, Switzerland

R. L. Hackney

Department of Physics and Astronomy, Western Kentucky University, Bowling Green, Kentucky 42101

C. Jordan

Department of Theoretical Physics, 1 Keble Road, Oxford, UK

C. G. Perola

Istituto di Fisica dell'Università di Milano, Milan, Italy

R. C. Roeder

Scarborough College, University of Toronto, Canada M5S 1A1

M. Schmidt

Hale Observatories, California Institute of Technology, Carnegie Institution of Washington, Washington 20005

During the commissioning phase of IUE several extragalactic objects were observed spectrally at low dispersion in the UV range $\lambda\lambda 1150$ – 3200 : the Seyfert galaxies NGC4151 and NGC1068, the QSO 3C273, the BL Lacertae object B2 1101+38, the giant elliptical galaxy M87 and the spiral galaxy M81. The results obtained are presented and a preliminary analysis given for all six objects, discussing the continuous spectrum, extinction, emission line spectrum and absorption line spectrum, where possible for each case. Several new or confirmatory astrophysical results are obtained.

DURING the commissioning high priority target phase of IUE several classes of extragalactic object were observed at low dispersion with both the short- and long-wavelength cameras. Although the instrument was not fully optimised, scientifically useful spectra were obtained for the Seyfert galaxies NGC4151 and NGC1068, the QSO 3C273, the BL Lacertae object B2 1101+38, the giant elliptical galaxy M87 and the spiral galaxy M81. A journal of observations is given in Table 1, which includes some related optical photometry (performed by A.S.). Figure 1 gives photographic reproductions of the short wavelength spectra of four of the objects observed. All six objects are discussed separately in following sections, and are preceded by a description of the adopted calibration, which differs in detail from that used for most of the other high priority targets discussed in the accompanying articles.

Absolute calibration

The initial absolute calibration of the scientific instrument is discussed by Boggess *et al.*¹. This is based on the OAOII scale² which differs from all other recent calibrations in the region between 1,350 and 1,700 Å, having a maximum departure of about 30% at $\lambda 1550$ (refs 3, 4). Extragalactic data reduced with the absolute calibration based on the OAOII scale show a hump

in the continuum around $\lambda 1550$. The same hump has been noted in the flux distribution of hot stars observed by IUE when reduced to the OAOII flux scale^{5,1}.

We have carried out further work on the absolute calibration because this is critical for establishing the slopes of the non-thermal continua. First, we linked the calibration mainly to the TD-1 (S2/68) fluxes, supplemented by ANS data. Second, we used additional objects to HD60753, the one used by Boggess *et al.*¹. This was accomplished without additional observations as some of the high priority objects, observed during the same period as the extragalactic objects presented here, are excellent calibration standards themselves. We used the following as additional calibration objects: BD+75 325, which was compared with OAOII data⁶, and data from ANS and TD1 S2/68 (P. R. Wesselius and D. J. Carnochan respectively, personal communications); HZ43, for which ANS data are available⁹; and HD137389, which has been observed by the S2/68 instrument¹⁰. The calibration for the stars listed by Boggess *et al.*¹ was converted to the S2/68 scale using the correction factors given by Beekmans³.

For the SWR camera, calibration curves obtained from the individual stars were averaged, giving double weight to the results for BD+75 325 and HZ43. The mean calibration curve so obtained was used for the reduction of all SWR extragalactic images. When compared with that given by Boggess *et al.*¹ this calibration curve shows two differences: the sensitivity minimum around $\lambda 1550$ is less pronounced, reflecting the difference between the OAOII and S2/68 or ANS flux scale (our continuum fluxes no longer have the $\lambda 1550$ hump), and the sensitivity at all wavelengths is lower. The extent of the second is comparable with the difference found by Boggess *et al.*¹ between the sensitivities obtained, after camera reoptimisation, for μ Col and HD60753. The four stars analysed for the SWR calibration all give roughly the same wavelength dependence for the sensitivity but the individual stars give absolute calibrations different by up to $\pm 20\%$ from the mean curve. Thus, the slope of a spectrum is well defined, but the overall flux level could be in error by $\pm 20\%$, in addition to any errors in the OAOII, S2/68

Table 1 Journal of observations

Date (1978)	Object	Camera,* band	Exposure (min)	Comments
11 February	NGC4151	SWR	60	Partly out of aperture; poor focus
11 February	NGC4151	LWP	120	
17 February	M87	LWP	180	
22 February	M81	SWR	120	
22 February	NGC1068	SWR	180	Good focus
27 February	3C273	SWR	180	
27 February	3C273	LWR	150	
27 February	3C273	SWR	60	
28 February	B2 1101+38	LWR	77	High radiation level
28 February	B2 1101+38	LWR	100	
28 February	B2 1101+38	SWR	240	
28 February	NGC4151	SWR	120	Good focus
28 February	B2 1101+38	UBVI		
7 March	NGC4151	UBV		
7 March	3C273	UBV		

*IUE ($10'' \times 2''$ aperture, low dispersion):

SWR (short wavelength redundant camera)

LWR (long wavelength redundant camera)

LWP (long wavelength prime camera)

Optical Photometry (Rosemary Hill Observatory, University of Florida):

UBVI

UBV.

or ANS calibrations and the errors involved in the transformation of the IUE data to other calibration systems (the latter is estimated to be $\sim 10\%$ for well-exposed spectra¹).

For the long wavelength cameras the high priority stellar targets permitted only a rough check for $2,000 \leq \lambda \leq 2,540$. For the accuracy of the measurements the results agree with the standard calibration¹, and this was used for the reduction of the objects in this paper.

Finally, we note that the 20% error mentioned above applies to both the short- and long-wavelength cameras, and the present results for the interstellar absorption maximum near $\lambda 2200$, which has to be measured on both the long and the short wavelength spectra for objects of low redshift, and is also subject to extra noise and uncertain end effects, should be treated as preliminary only.

NGC4151

Comprehensive reviews of observational results and model computations for Seyfert galaxies are given by Weedman¹¹ and Collin-Souffrin⁶⁶. The well-known Seyfert galaxy NGC4151 is the brightest of the class with broad permitted emission lines (Class 1 of Khachikyan and Weedman¹²) in the northern skies and has been closely studied at all available wavelengths. The permitted lines also contain a distinct narrow component with the same profile as the forbidden lines (which are always relatively narrow in Seyfert galaxies) and NGC4151 is better classified 1.5. Its nucleus contains a compact source which is variable at X-ray, UV, optical and IR wavelengths^{13-17,49}. It is also a γ -ray source¹⁸ and a weak source of radio emission¹⁹, although the radio emitting volume may be different from the region producing the optical continuum and broad emission lines²⁰. Recently an UV spectrum of NGC4151 has been obtained by Davidsen and Hartig with a rocket-borne telescope⁴⁶.

The optical spectrum of NGC4151 has been discussed by several authors²¹⁻²⁵. The narrow-line region has temperature $T = (11,750 \pm 800)\text{K}$, and electron density $N_e = 2 \times 10^4 \text{ cm}^{-3}$ in a higher ionisation region and $N_e = 5 \times 10^3 \text{ cm}^{-3}$ in a lower ionisation zone defined by [O I], [O II] and [S II], which will be excited at the outer portions of the gas²⁴. The relative intensities of the forbidden lines can be well matched by photo-ionisation models in which the gas is ionised by a non-thermal UV continuum, and probably this is also the main energy source for the broad-line region^{26,27}. The broad-line region is characterised by the permitted emission lines of H I, He I, He II and Fe II, and must have $N_e \geq 10^8 \text{ cm}^{-3}$ to collisionally de-excite the forbidden lines which are absent^{28,24}. High column densities are also indicated by the Balmer decrement, which can be understood

only if the higher Balmer lines are optically thick^{22,23,39}. As is commonly observed in such objects, the broad components of the Balmer lines in NGC4151 are asymmetric and extend over several thousand km s^{-1} ; the He II $\lambda 4686$ profile is similar, but He I $\lambda \lambda 5876$ and 7065 are much broader²⁴. The asymmetries and different linewidths rule out electron scattering as the dominant broadening mechanism, and invalidate the model for the line emitting region developed by Ptak and Stoner^{29,24}. There is now little doubt that the broadening is chiefly due to mass motions of the ionised gas in such active nuclei^{27,30,31}.

A striking feature of the optical spectrum of NGC4151 is the presence of non-stellar Balmer and He I $\lambda 3889$ absorption features which are blueshifted with respect to the emission lines and variable in time^{32,33,25,140}. These features are characteristically different from the absorption lines generally present in QSO spectra, which arise only from ground-state transitions and have never been seen to vary. Galactic interstellar Ca II absorption is also present in the spectrum of NGC4151 (refs 24, 25, 140).

NGC4151 was observed by IUE with both the short and long wavelength cameras (Table 1). The better exposed of the two short wavelength spectra (28 February) is shown in Fig. 2. Emission features (28 February for short wavelength, 11 February for long) are listed in Table 2 and the absorption features (28 February) in Table 3. Optical spectra taken soon after (8 March) (J. McClintock, personal communication) show that the visual absorption lines were particularly strong. For the absorption lines arising from levels of the ground state there is probably a significant galactic interstellar contribution in some cases (the heliocentric velocity of NGC4151 determined from the emission lines is about 950 km s^{-1} (ref. 32), and the absorption lines are blueshifted relative to NGC4151 by a comparable amount^{32,140}, bringing them to near-zero velocity in the observed frame).

The continuous spectrum of NGC4151 obtained on 28 February, with measurements averaged over bands 20–100 Å wide avoiding the emission and absorption lines, is shown in Fig. 3. We include our UBVI photometry (7 March), corrected for the emission lines and Balmer-continuum emission (the latter from the expression given by Stein and Weedman³⁴) and reduced to the same effective aperture as the IUE observations ($10'' \times 20''$) from the data of Penston *et al.*¹⁶. We also show the continuum data of Davidsen and Hartig⁴⁶, obtained on 10 February.

The observed optical continuum of an active galaxy nucleus is made up of two parts: the featureless non-thermal or 'power law' continuum unique to the active nucleus, and the integrated galaxy continuum (containing absorption lines) due to stars in or near the nucleus. (The term 'non-thermal' is used rather loosely; 'non-stellar' might be better in some cases where a thermal source, possibly complex, cannot be ruled out.) When the continuum shows variability the spectral slope generally also varies due to the differing contributions of the non-thermal variable component and the stellar component, although the slope of the non-thermal component probably does not change¹⁴. Assuming a power-law spectrum of form $F_\nu \propto \nu^{-\alpha}$ we

Fig. 1 Photographic reproductions of unprocessed short wavelength spectra of four of the extragalactic objects: first, the Seyfert 1.5 galaxy NGC4151; second the Seyfert 2 galaxy NGC1068; third the QSO 3C273; fourth the BL Lac object B2 1101+38 (Markarian 421). Wavelength increases to the right. The large emission feature at the left of each spectrum is geocoronal Ly α which fills the $10'' \times 20''$ entrance aperture.

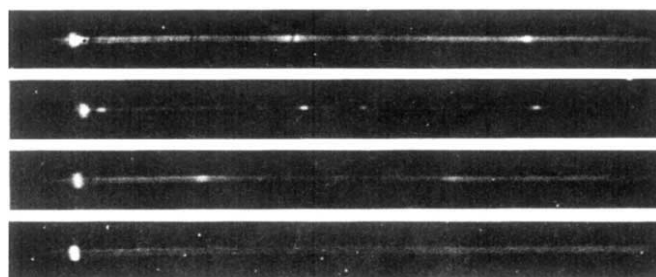


Table 2 Emission lines in the spectrum of NGC4151 ($\lambda < 2,000$, 28 February, spectrum; $\lambda > 2,000$, 11 February spectrum)

Identification	λ_0 (Å)	W_{λ_0} (Å)	F (10^{-12} erg cm $^{-2}$ s $^{-1}$)	Comments	
Ly α	1,216	103 $\pm$ 45	29 $\pm$ 13	Estimated from less-contaminated red half Uncertain because of presence of absorption lines	
O I	1,305	—	—		
Si IV + O IV]	1,400	—	—		
N IV	1,486	3†	0.7†		
C IV	1,550	110†	24†	Approximately corrected for C IV absorption, uncertain Possibly includes wing of C IV	
[Ne v]	1,575				
[Ne IV]	1,602				
He II	1,640	11	2.3		
O III]	1,663	5	0.9		
N III]	1,749	2	0.3	$W_{\lambda_0} \sim 7$ Å for Si III alone	
Si III]	1,892	40	6.1		
C III]	1,909				
[O III] + C II]	2,324				
[Ne IV]	2,424				
?	2,450	10	1.6	Corrected for absorption Too strong for O III alone Reseau present	
[O II]	2,470				
?	2,500*	2†	0.4†		
[Fe XI]	2,650				
Mg II	2,800	48†	7.8†		
O III + [Fe IV]?	2,837	5	0.7		
O III	3,047	≥ 3	≥ 0.4		
O III	3,133	8	1.0		

* Wavelength measured relative to other known lines.

† Uncertain.

obtain $\alpha = 1.3$ as the mean over the observed optical and UV range, excluding possible absorption attributed to the extinction feature near $\lambda 2200$. Contamination of the non-thermal continuum by light from the galactic disk of NGC4151 occurs to differing degrees in the optical and UV regions. Comparison with the results of Wu and Weedman¹⁴ shows that for the IUE entrance aperture and the aperture of 36" diameter used by Davidson and Hartig⁴⁶ contamination is negligible in the UV. We obtain a shallower slope, $\alpha \sim 0.8$, for $1,300 < \lambda < 1,800$, in agreement with Davidsen and Hartig⁴⁶. However, our continuum flux here is about a factor 2.5 higher than theirs, obtained 18 d earlier. Furthermore, our long wavelength continuum fluxes obtained on 11 February (only one day later than the short wavelength observations of Davidsen and Hartig) and our UVB photometry of 7 March, are both consistent with our short wavelength continuum, but even our 11 February data are out of line with Davidsen and Hartig's result. The fluxes of the emission features observed by both IUE and Davidsen and Hartig (excluding the contaminated Ly α line) show no

significant change between the extreme dates. Furthermore, we observe no change in equivalent width of the C III] + Si III] emission between our two IUE short wavelength observations, 17 d apart. We conclude that there was no significant continuum change over the period of the IUE observations but there may have been a large and extremely rapid enhancement between 10 February (Davidsen and Hartig) and 11 February (IUE). On the other hand, an error in defining the zero point of the flux scale in Davidsen and Hartig's data could be an alternative explanation. A gross calibration error seems to be ruled out by the generally good agreement for the emission line fluxes.

Even with no correction for reddening, extrapolation of our far UV continuum gives sufficient ionising radiation to account for the strength of the Balmer emission lines by recombination in a gas that nearly covers the source and is optically thick to the Lyman continuum²⁴. This is contrary to the interpretation of Wu and Weedman¹⁴ using ANS data. We also find that our spectrum extrapolates reasonably well to the X-ray observations, again contrary to Wu and Weedman¹⁴. The differences between our results and those of Wu and Weedman seem to be due to the way they correct for emission lines in the ANS photometric channels. Comparison with Fig. 2 and Table 2 shows a substantial emission line contribution to the $\lambda 2500$ channel and to their $\lambda 1550$ 'continuum' channel. If we use their definition of 'continuum' fluxes in the various channels and then calculate IUE 'continuum' fluxes we obtain the same spectral slope as they derive.

An important constituent of the active nuclei of Seyfert galaxies is dust, which evidently is intimately mixed with the ionised gas in these objects. The ratio of the [S II] $\lambda \lambda 4072, 10320$ lines³⁵ gives a colour excess $E(\lambda 4072 - \lambda 10320) = 0.82$. Using $E(\lambda 4072 - \lambda 10320) = 3.37 E(B - V)$ from Johnson and Borgman³⁶ we obtain $E(B - V) = 0.24$. Penston and Fosbury⁸⁴ give $E(B - V) = 0.6$ from the Balmer decrement of the narrow lines. A characteristic dust signature in the UV is a deficiency of flux near $\lambda 2200$. We estimate $W_{\lambda}(2200) \sim 30 \text{ Å}$ from our data (this is uncertain by a factor ~ 2) implying $E(B - V) \sim 0.06$ if the galactic extinction law applies³⁷, in (fortuitous) agreement with the value obtained by Wu and Weedman¹⁴. That this is much less than the value derived from the [S II] lines or the Balmer lines may not be a serious difficulty because reddening is not necessarily the same in the continuum, broad- and narrow-line emitting regions; but there is probably some contribution from dust in our Galaxy. Applying a reddening correction (using $E(B - V) = 0.06$ and the extinction law given by Code *et al.*³⁸) to our far UV continuum slope gives $\alpha \approx 0.7$; applying it to all the data in Fig. 3 gives $\alpha \approx 1.0$.

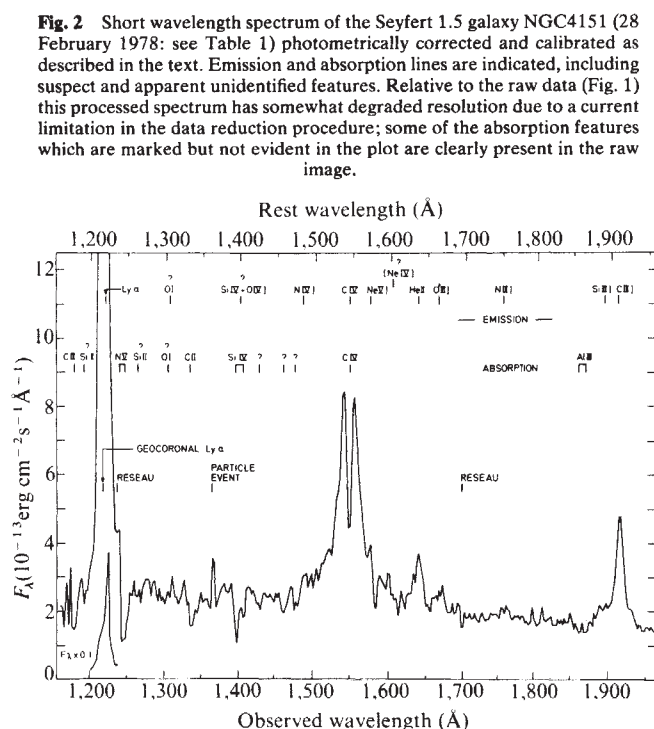


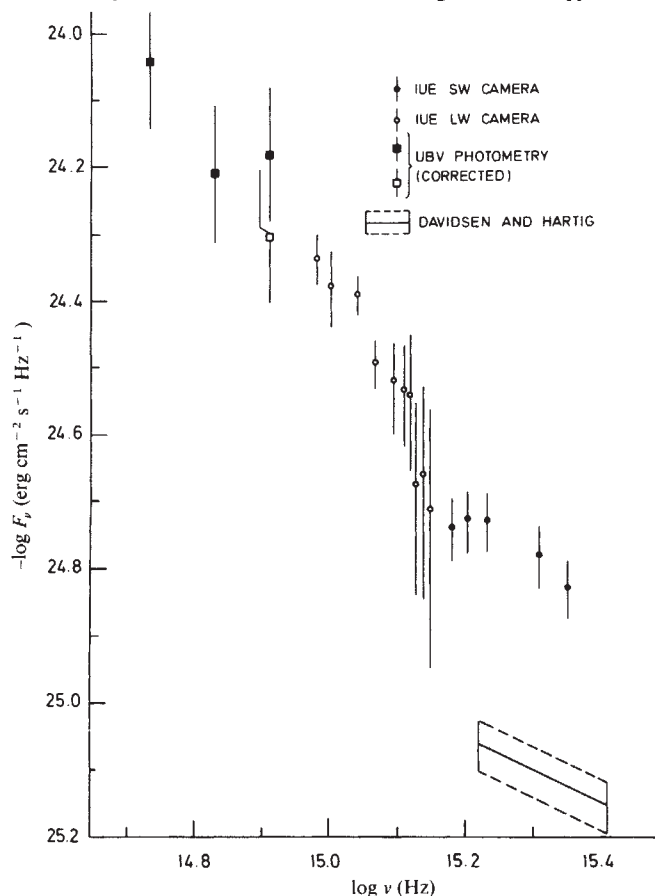
Fig. 2 Short wavelength spectrum of the Seyfert 1.5 galaxy NGC4151 (28 February 1978; see Table 1) photometrically corrected and calibrated as described in the text. Emission and absorption lines are indicated, including suspect and apparent unidentified features. Relative to the raw data (Fig. 1) this processed spectrum has somewhat degraded resolution due to a current limitation in the data reduction procedure; some of the absorption features which are marked but not evident in the plot are clearly present in the raw image.

The different degrees of reddening obtained for the continuum and forbidden line regions is interesting because it shows that our view of the continuum source is relatively unobscured by line-emitting gas, which therefore does not totally surround the continuum source. However, a covering factor as little as 0.7 is consistent with the observed strength of $H\beta$ and our far UV continuum slope. A model of the ionised gas in Seyfert nuclei deduced from observational data is given by Osterbrock²⁷. We return to this later.

The spectrum of NGC4151 shows strong emission lines of $Ly\alpha$, C IV λ 1550, He II λ 1640, C III] λ 1909, [O III] + C II] λ 2324, [Ne IV] λ 2424 and Mg II λ 2800, and several weaker lines (Fig. 2 and Table 2). There is little evidence of N V λ 1240, Si II λ 1263, O I λ 1305, C II λ 1335 or Si IV + O IV] λ 1400 in emission, but for these the spectrum is confused by definite or possible absorption features. Nevertheless, N V λ 1240, if present, is very weak. The presence of [Fe XI] λ 2650 is consistent with the detection of the coronal line [Fe XI] λ 7892 in this object⁴⁵.

The $Ly\alpha$ profile is affected by the geocoronal emission, a fiducial mark (reseau) and $Ly\alpha$ and N V absorption, but it seems to have comparable width to the Balmer lines²⁴. The C IV profile also is strongly affected by absorption, which makes it difficult to determine the FWHM, but it seems peculiarly broad with very extensive wings. Accounting for contributions from N IV] λ 1486, [Ne IV] λ 1602 and possibly [Ne V] λ 1575 still leaves a very broad asymmetric line of full width at zero intensity $\approx 15,000 \text{ km s}^{-1}$, which is comparable with the broad He I lines observed in the visible²⁴ and suggests a common origin. Alternatively, there may be a strong contribution from Fe III to the C IV profile, as contended by Davidsen and Hartig⁴⁶, but

Fig. 3 NGC4151 continuum fluxes (28 February 1978 for SW, 11 February 1978 for LW: see Table 1). The IUE data are averaged over bands 20–100 Å avoiding the emission lines. Our UVB photometry (6 March 1978) reduced to the IUE observing aperture, also are corrected for emission lines (■) and additionally for Balmer continuum emission (□). The continuum data of Davidsen and Hartig⁴⁶, obtained on 10 February 1978, are included. The error bars indicate 1σ statistical uncertainties in signal only, and do not include possible calibration errors. No reddening correction is applied.



according to the results of Viotti⁴⁷ Fe III emission around λ 1550 should be accompanied by even stronger Fe III emission between $\lambda\lambda$ 1840 and 2020, which is absent in our data. In contrast, C III] λ 1909 is relatively narrow, but a broad component of comparable width to that of the Balmer lines is clearly present, blended with Si III] λ 1892 in the blue wing. This sets an upper limit of $N_e \leq 10^{10} \text{ cm}^{-3}$ to the broad-line region.

Netzer³⁹ has shown that a steep Balmer decrement can be produced if $N_e > 10^8 \text{ cm}^{-3}$, in explanation of observed inconsistencies when attempting to fit extinction-corrected Balmer line strengths for QSO and Seyfert galaxies to values predicted from radiative recombination theory. Baldwin⁴⁰ has pointed out a more extreme difficulty for the intensity ratio of $Ly\alpha$ to $H\beta$ in QSOs, which is directly observed or deduced to be ≈ 3 (refs 40–42). Recombination theory predicts values 22–40 depending on the extent that $Ly\alpha$ is collisionally excited. While selective extinction by dust could reduce the relative intensity of $Ly\alpha$, particularly if it has a large optical depth compared to $H\beta$ and is multiply scattered, it is curious that $I(Ly\alpha)/I(H\beta)$ is so remarkably constant among cases having substantially different apparent reddening. This matter has been further discussed recently by Zirin⁵⁰. For the Seyfert galaxy NGC4151, also, we find $I(Ly\alpha)/I(H\beta) = 4.4 \pm 2$ using $I(H\beta) = 6.6 \times 10^{-12} \text{ erg cm}^{-2} \text{ s}^{-1}$ (refs 21, 22, 24). Assuming that $E(B - V) = 0.06$ holds not only for the continuum but also for the permitted line region (which emits most of the flux in the hydrogen lines), and making a simple application of the same extinction law as before, gives $I(Ly\alpha)/I(H\beta) = 6.3 \pm 3$. On the other hand, if $E(B - V) = 0.24$, as deduced for the narrow-line region from the [S II] lines, we obtain $I(Ly\alpha)/I(H\beta) = 18 \pm 8$, closer to the recombination value. Note that the presence of geocoronal $Ly\alpha$, $Ly\alpha$ absorption in the Galaxy and in NGC4151, and the camera fiducial mark in the red wing of the NGC4151 $Ly\alpha$ line makes the measurement of this line rather uncertain.

Baldwin⁴³ has found that for QSOs there exists an important correlation between $W_{\lambda_0}(\text{C IV})$, the rest frame equivalent width of the C IV λ 1549 emission doublet, and $L_{\nu}(1450)$, the continuum luminosity at a rest wavelength λ 1450 computed on the assumption that the redshift is a cosmological indicator of distance, suggesting that these objects can be used as standard candles. The relationship seems particularly tight if restricted only to the flat spectrum radio sources, as confirmed by Baldwin *et al.*⁴⁸. The similar study by Osmer⁴⁴ is in apparent conflict, showing $W_{\lambda_0}(\text{C IV})$ to be independent of $L_{\nu}(1475)$, but he covered a smaller luminosity range and his results are consistent with Baldwin's more widely ranging data. Comparing our (rather uncertain) value of $W_{\lambda_0}(\text{C IV})$ for NGC4151 with

Table 3 Absorption lines in the spectrum of NGC4151 ($\lambda < 2,000$, 28 February spectrum; $\lambda > 2,000$, 11 February spectrum)

Identification	λ_0 (Å)	W_{λ_0} (Å)	Comments
C III*	1,176	4†	
Si II	1,193	1†	
N V	1,239, 1,243	6†	
Si II	1,260	—	Weak
Si II*	1,265	—	Weak
O I	1,302	—	Possible
C II, C II*	1,335, 1,336	3†	Seems diffuse
Si IV	1,394	3.5†	
Si IV	1,403	1.5†	
?	1,425*	1†	Possible
?	1,456*	1†	Possible
?	1,473*	—	Possible
C IV	1,548, 1,551	>3	In strong emission line
Al III	1,855	1†	
Al III	1,863	1†	
?	1,963*	1†	Possible
Fe II	2,587–2,632	4†	Broad shallow feature
Mg II	2,796, 2,803	$2 \leq W_{\lambda_0} \leq 6$	Less strong emission line; probably partly galactic

*Wavelength measured relative to other known lines.

†The unknown contribution of the emission lines to the local 'continuum' for the absorption lines makes the values of equivalent width for these uncertain.

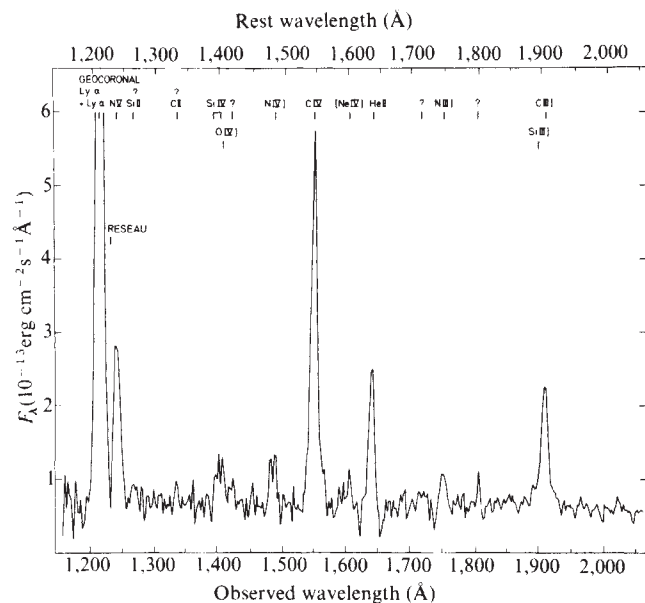


Fig. 4 Short wavelength spectrum of the Seyfert 2 galaxy NGC1068, photometrically corrected and calibrated. Emission lines are indicated, including suspect and apparent unidentified features. The Ly α line is saturated, and contaminated by geocoronal Ly α . The almost complete separation of Ly α and N V is real, not due to the fortuitous presence of the reseau (see Fig. 1).

Baldwin's relationship gives an expected luminosity $\log L_{\nu}(1450) = 30.1$, but we find the value 27.9, which is in serious conflict. Later we shall show the Seyfert galaxy NGC1068 and the BL Lac object B2 1101 + 38 (which is a flat spectrum radio source) also do not fit the relationship but the QSO 3C273 does fit well. We conclude that Baldwin's luminosity relationship at least is not generally applicable to active nuclei, and with the available evidence seems to segregate QSOs from Seyfert galaxies and (less surprisingly) BL Lac objects. However, excluding B2 1101 + 38, we note that over the large luminosity range now obtaining with the inclusion of NGC4151 and NGC1068 with the sample of QSOs, the data broadly conform with Osmer's⁴⁴ results which show no luminosity dependence for $W_{\lambda 0}$ (C IV).

From model computations Shields⁵¹ has shown that use of some of the weaker emission lines allows abundance determinations that are not strongly dependent on the details of the photoionisation models for active nuclei; in particular, these are: O III] λ 1663, N IV] λ 1486, N III] λ 1749 and C III] λ 1909. Baldwin and Netzer⁵² made a refined analysis and find some minor modifications to Shield's results. Using Shield's relationships we find the relative abundance N/C is 0.16 from N IV]/C IV and 0.13 from N III]/C III]. Baldwin and Netzer⁵² find N III]/C III] is ~ 1.5 larger in their analysis, and using their correction we obtain N/C = 0.09. For O/C and N/O we obtain 1.37 and 0.20 from use of O III]/C IV and N IV]/O III] respectively. Given the uncertainties in our line fluxes, these values are reasonably self-consistent and give relative abundances N/C/O which are comparable with solar values¹⁴¹.

Our two short wavelength spectra were obtained 17 d apart (Table 1). Although the 11 February exposure is somewhat out of focus and underexposed, there is clear absorption present at C IV λ 1550, Si IV λ 1394, 1403 and C II λ 1335, 1336. If N V λ 1240 is present it is rather weak. Our well-exposed and in-focus spectrum of 28 February shows the several lines listed in Table 3. N V is strong, C II seems weaker, and C IV stronger with the absorption more centrally placed in the emission line, than on 11 February. C II seems rather diffuse (28 February spectrum), contrasting with the sharp Si IV and Si II lines. Time-varying structural differences between the He I λ 3888 and Balmer line absorption features have been observed in the optical spectrum^{141,32}. In the spectrum obtained by Davidson and Hartig⁴⁶ one day earlier than our first spectrum, N V, Si IV

and C IV seem absent, Si II and C II are strong, and there may be other features. Their resolution is lower than ours, and C IV, for example, may be present but lost in the strong emission profile. However, at face value there is evidence that the ionisation of the absorption system has changed greatly over the period of 18 d between the observations of Davidson and Hartig and our good 28 February spectrum, and to some degree in the one day between their observation and ours of 11 February. Some change in line strength is evident over 17 d from the IUE data alone. The dramatic increase in ionisation apparent from the time of Davidson and Hartig's observations is accompanied by a large increase in the far UV continuum flux discussed earlier, if real.

The presence of the C III* λ 1176 absorption line, which arises from the metastable $2p^3P^0$ state, is consistent with the excitation of the optical spectrum^{25,32,140}. Si II* λ 1265 arises from an excited fine structure level, and seems comparable in strength to the ground state line λ 1260, suggesting that the electron density in the absorbing region is $N_e > 10^4 \text{ cm}^{-3}$ if the level is populated by collisions at $T_e \approx 10^4 \text{ K}$. However, the Si II lines are only marginally detected.

At our relatively low resolution we cannot determine the true depth of the absorption lines, but it is clear from the observed depths of the various strong absorption features that not only the continuum but also a substantial fraction of the broad-line-emitting region is covered by the absorber. Thus, the absorbing region cannot be a component (filament) of the broad-line-emitting region but must be more extensive and outside it. A similar conclusion was reached by Anderson³² from the optical data.

NGC1068

NGC1068 is the archetype of class 2 Seyfert galaxies, and has been well studied at optical, IR and radio wavelengths^{16,20,54-63,68}; it shows no convincing variability. It has an IR nucleus $\sim 100 \text{ pc}$ in diameter, which is inconsistent with a single non-thermal source, but is consistent with emission from heated dust⁵⁸. Emission features due to H₂ are present in the IR spectrum⁶², and the absorption feature near $9.5 \mu\text{m}$ which has been attributed to silicates is strong⁶⁴. It has somewhat broader lines than other Seyfert 2 galaxies, with FWHM about $1,200 \text{ km s}^{-1}$, but they are still considerably narrower than in Seyfert 1 galaxies. The lines have an asymmetric form with much velocity structure, and spatial structure discernible on a scale of $1''$ or 70 pc (ref. 65) and extending over $\sim 8''$ (ref. 59). The line emitting region is physically complicated, having regions of different density, filling factor and degree of shadowing from the continuum source^{57,56}. NGC1068 is another of several Seyfert galaxies found to emit the coronal line [Fe XI] λ 7892 (ref. 45).

NGC1068 was observed by IUE only with the short wavelength camera. The spectrum is shown in Fig. 4. Only emission features are definitely observed, and these are listed in Table 4. The spectrum was over-exposed in Ly α , and this line

Table 4 Emission lines in the spectrum of NGC1068

Identification	λ_0 (Å)	$W_{\lambda 0}$ (Å)	F ($10^{-13} \text{ erg cm}^{-2} \text{ s}^{-1}$)	Comments
N V	1,240	43	29	
Si II	1,263	4†	3†	Possible
C II	1,335	3†	2†	Possible
Si IV + O IV]	1,400	13	9	
?	1,420*	4†	3†	Possible
N IV]	1,486	11	7	
C IV	1,550	90	59	
[Ne IV]	1,602	5†	3†	
He II	1,640	28	18	
?	1,720*	7	4	Probable, maybe N IV
N III]	1,749	6	4	
?	1,800*	—	—	Possible
Si III]	1,892	40	25	
C III]	1,909			

*Wavelength measured relative to other known lines

†Uncertain.

also is contaminated by the geocoronal emission, so we have no reliable flux estimate. The almost complete separation of Ly α and N V λ 1240 is real, not due to the fortuitous presence of the reseau (see Fig. 1).

Continuum fluxes, mostly averaged over about 40 Å and avoiding the emission lines, are shown in Fig. 5, together with part of Wampler's⁵⁵ spectrophotometry also avoiding emission lines and reduced to the IUE observing aperture with the aid of the data of Penston *et al.*¹⁶. We find $\alpha = 2.4$ fitted to the usual power law as a mean over the whole optical and UV region included in Fig. 5. With our aperture, the visual flux of the nuclear source is much weaker than that of the underlying galaxy^{16,69} but in the far UV (IUE) region, where contamination from the galactic component is negligible, we obtain the important result that the continuum clearly is a power-law spectrum, with $\alpha \approx 1.8$ (not accounting for reddening); there is no unambiguous evidence of a power-law continuum in the optical region. We have no direct estimate of reddening from the IUE data (the region containing the λ 2200 absorption feature was not observed) but if we apply a correction based on the value $E(B-V) = 0.5$ derived from the [S II] lines and the Balmer decrement^{55,67} the resulting spectrum is highly inconsistent with a power law continuum, suggesting that $E(B-V)$ for the continuum is much smaller (≤ 0.15) and implying $\alpha \geq 1.4$. However, a flatter continuum than observed may be needed to explain the presence of [Fe XI] (ref. 45) and the observed value of $I(\text{He}^+)/I(\text{H}^0)$ (ref. 57).

The different degrees of reddening deduced for the continuum and emission line regions again shows that the continuum is not completely surrounded by line-emitting gas. This is consistent with optical and radio data^{57,61}.

The spectrum of NGC1068 shows strong, relatively narrow, emission lines of Ly α , N V λ 1240, C IV λ 1550, He II λ 1640 and C III] λ 1908 (Fig. 4 and Table 4). Several weaker lines are definitely present, including Si IV + O IV] λ 1400, N IV] λ 1486 and N III] λ 1749.

Comparing our value of W_{λ_0} (C IV) for NGC1068 with Baldwin's⁴³ empirical relationship coupling this quantity with

Fig. 5 NGC1068 continuum fluxes. The IUE data are averaged over about 40 Å bands. Wampler's⁵⁵ spectrophotometry is reduced to the IUE observing aperture. Both sets of data are chosen to avoid emission lines. The error bars indicate 1σ statistical uncertainties only. No reddening correction is applied.

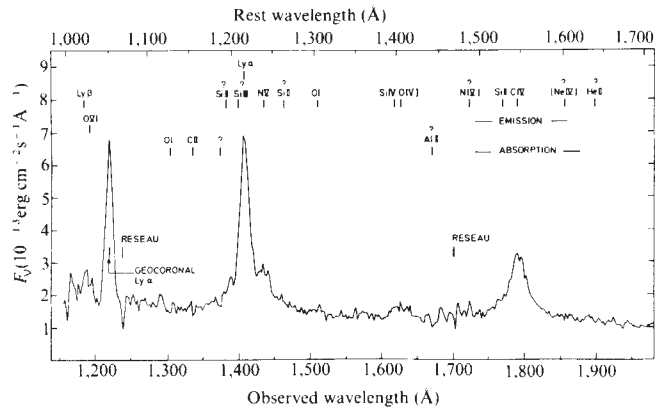
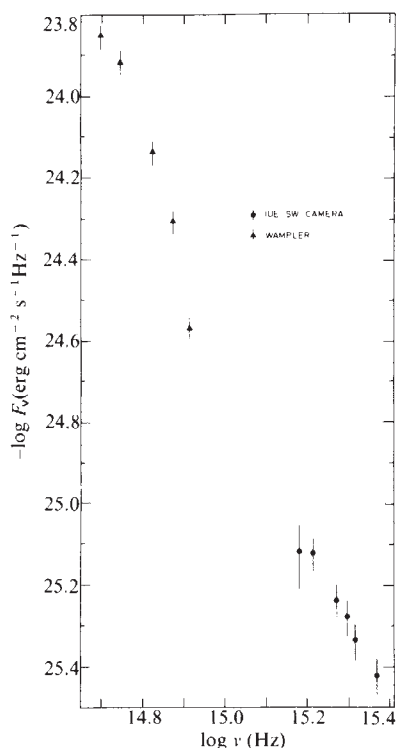


Fig. 6. Short wavelength spectrum of the QSO 3C273, photometrically corrected and calibrated. Emission and absorption lines are indicated (see related comments in the caption to Fig. 2).

$L_v(1450)$, as we have done for NGC4151, gives an expected luminosity $\log L_v(1450) = 30.1$, but we again find a much lower value, 27.4. Thus, we find serious (and similar) departures from Baldwin's relationship for both the Seyfert galaxies studied here.

Calculating the relative abundance N/C from the work of Shields⁵¹ and Baldwin and Netzer⁵², as before, we find the value ~ 1.0 , nearly six times the solar value¹⁴¹. Shields and Oke⁵⁷ find a fourfold over-abundance for N/O from optical data for this object.

3C273

Like the other bright extragalactic objects discussed here the QSO 3C273 ($z_{\text{em}} = 0.158$) has been very well studied at all available wavelengths^{70-80,84,58}. Far UV observations preceding the present work were made by Wu⁸ and Davidsen *et al.*⁴². Apart from being the brightest QSO, 3C273 is also one of the most luminous. It is only moderately variable in the radio, IR, optical and X-ray regions^{58,79}. VLBI radio measurements indicate a double component apparently separating at a projected linear velocity $\sim 4c$ (ref. 80). 3C273 is among comparatively few QSOs which show the forbidden lines to be very weak, indicating that there is very little low density gas in the line-emitting region⁷⁵. It is also rare in showing strong Fe II emission^{72-75,81}.

3C273 was observed by IUE with both spectrograph cameras, as detailed in Table 1. One of the two short wavelength spectra is shown in Fig. 6. Emission features are listed in Table 5. Weak absorption features are also present, listed in Table 6, but are not easy to discern in Fig. 6 because at this time the telescope focus was rather soft (in Fig. 1 compare the raw spectrum of 3C273 with that of NGC4151 which was in sharp focus, although trailed). We identify all but one doubtful line as arising in interstellar gas in the Galaxy. Corresponding hydrogen absorption at Ly α is not observed, but this region is masked by geocoronal Ly α emission as a consequence of using the large entrance aperture.

The continuous spectrum of 3C273, with measurements averaged over bands 20–50 Å wide avoiding the emission lines, is shown in Fig. 7. We include our UVB photometry, and Wu's⁸ ANS photometry, both corrected for emission lines, and the continuum data of Davidsen *et al.*⁴². Assuming the usual power law we find $\alpha \approx 1.0$ for the slope of the continuum over our whole observed UV (IUE) range, uncorrected for possible reddening. In optical range, as is well known^{42,75}, the spectrum is much flatter and contains strong Balmer continuum emission (in the U- and B-channel). Davidsen *et al.*⁴² argue that for 3C273 there is no internal reddening and for galactic reddening they deduce $E(B-V) = 0.01$; but Wu⁸ estimates $E(B-V) = 0.07$ for galactic reddening, and from an apparent absorption at redshifted λ 2200 obtains $E(B-V) = 0.13$ for internal reddening. At least for the line emitting region, internal reddening is

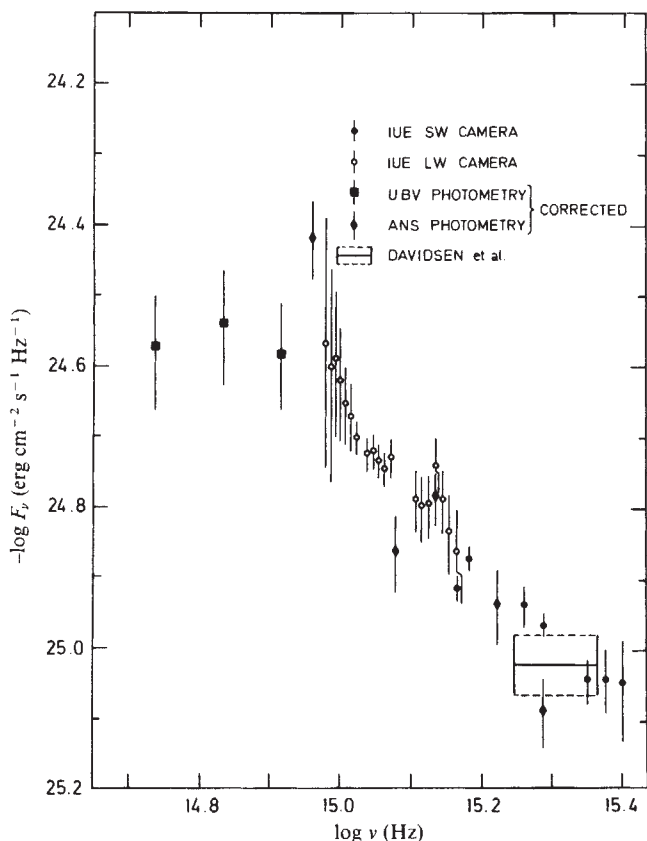


Fig. 7 3C273 continuum fluxes. The IUE data are averaged over bands 20–50 Å wide avoiding the emission lines. Our UBV photometry and the ANS photometry⁸ also are corrected for emission lines. The continuum data of Davidsen *et al.*⁴² are included. No reddening correction is applied. The error bars indicate 1σ statistical uncertainties only. Although the various observations were not made concurrently, we may rely on the general character of these continuum data as 3C273 is known not to be highly variable.

suggested by the observed Balmer decrement, which is steeper than the predicted recombination values and perhaps indicates $E(B-V) = 0.69$ (ref. 75), although reddening alone does not seem a sufficient explanation⁷¹. However, reddening is not confirmed by the ratio $\text{Pa}/\text{H}\alpha$ (ref. 83), nor the helium line strengths⁷⁵, showing that the situation is more complicated than supposed^{75,73,39}. The IUE spectrum is somewhat contaminated near the redshifted position of the $\lambda 2200$ dust feature, but there is no sign at all of broad absorption at this wavelength, and on this basis we conclude that internal reddening is small for the 3C273 continuum.

3C273 shows strong line emission at $\text{Ly}\alpha$ and C IV $\lambda 1550$, moderate emission at $\text{Ly}\beta + \text{OVI } \lambda 1035$, N V $\lambda 1240$, Si IV + O IV $\lambda 1400$, C III] $\lambda 1909$ and Mg II $\lambda 2800$, and has some definite or possible weaker lines, as shown in Fig. 6 and Table 5. There is no obvious difference between the UV emission line spectrum of 3C273 and those of high redshift QSOs observed optically^{52,85,86,44}, except that C III] $\lambda 1909$ seems weaker than average, suggesting a relatively high density in the line emitting region. The ratio of line strengths $I(\text{N V } \lambda 1240)/I(\text{Ly})$, ~ 0.09 , is within the range observed for high redshift QSOs, which show a wide variation in this ratio^{44,52,86,87}. Once more we find that the ratio $I(\text{Ly}\alpha)/I(\text{H}\beta)$ is much lower than predicted by recombination theory; combining our measurement of the $\text{Ly}\alpha$ flux with the value obtained by Davidsen *et al.*⁴² and comparing this with previous measurements of the $\text{H}\beta$ flux^{75,84} gives $I(\text{Ly}\alpha)/I(\text{H}\beta) = 5.5 \pm 2$. Again we note the remarkable constancy of the observed ratio. Assuming no reddening in the line-emitting region but applying Wu's⁸ galactic reddening correction raises the ratio to 7 ± 3 ; using the smaller correction of Davidsen *et al.*⁴² gives the value 6 ± 2 .

Comparing our value of $W_{\lambda 0}$ (C IV) for 3C273 with Baldwin's⁴³ luminosity relationship, as before, gives an expected

luminosity $\log F_{\nu}(1450) = 30.6$. Our observed value is 31.0, in good agreement. As already stated, with the available evidence the range of validity of Baldwin's relationship among active nuclei seems to be restricted to QSOs. Baldwin's result is not expected from simple photoionisation models which naturally lead to a constancy of equivalent width with luminosity, as found by Osmer⁴⁴: for a given spectral index the continuum flux is an indicator of the energy available for heating the gas and the C IV line is a good measure of the energy emitted by the gas; hence the equivalent width, that is, the line strength expressed in terms of the continuum, should be independent of luminosity. However, Osmer⁴⁴ does find a significant luminosity dependence for $W_{\lambda 0}(\text{Ly}\alpha)$, also noted by Baldwin⁴³, to an extent not expected in the photoionisation model.

The optical Fe II emission lines in 3C273 all arise from transitions between odd-parity levels at $\sim 5\text{eV}$ and even-parity levels at $\sim 3\text{eV}$ (refs 82, 113). The transitions from these 3eV levels down to the lower-lying levels of the ground state are forbidden. These [Fe II] transitions are not observed in 3C273, indicating that the electron density $N_e > 10^7 \text{cm}^{-3}$ in the Fe II emission region. The most likely mechanisms for producing the observed Fe II emission are resonance fluorescence of continuum radiation¹¹³ and collisional excitation^{78,82,114,115}, the former being the one generally preferred although the correlations implied by this mechanism are not observed¹⁴². The strongest Fe II lines in 3C273 belong to multiplet 42 ($\lambda\lambda 4924, 5018$ and 5169). A test of resonance fluorescence is to observe the spectral region of the resonance lines in multiplet 3UV (around $\lambda 2340$) and look for absorption of continuum photons¹¹⁶. No such absorption is observed for 3C273. (Fe II absorption lines are present in NGC4151, which also shows weak optical Fe II emission, but these are most likely in part due to the unusual absorbing region observed in this object and in part galactic). Model calculations^{82,113} imply nearly complete conversion of the UV continuum to explain the strength of the optical Fe II emission, so it is surprising that no obvious absorption is present. On the other hand, collisional models predict strong emission in the resonance lines⁸², which also is not observed. In fact, the fluorescence models also predict strong resonance line emission, complementing the continuum absorption in the same lines. Depending on the geometry of the line-emitting region, fluorescence models could produce either net absorption or emission in the resonance lines, so the failure to observe either does not rule out the mechanism⁸². Again, collisional models imply a very large optical depth in the resonance lines⁸² and substantially all resonance photons may then be destroyed through the leak which gives the optical lines. The effect of increasing optical depth has been discussed by Jordan¹⁴³. But then resonance fluorescence of continuum radiation would also occur, so collisional excitation alone is not likely

Table 5 Emission lines in the spectrum of 3C273

Identification	λ_0 (Å)	$W_{\lambda 0}$ (Å)	F ($10^{-12} \text{ erg cm}^{-2} \text{ s}^{-1}$)	Comments
Lyβ	1,026			
O VI	1,035	5†	1.1†	
Si II	1,194	2†	0.3†	Probable
Si III	1,206	79	13.5	
Lyα	1,216			
N V	1,240	7	1.2	
Si II	1,263	—	—	Probable
O I	1,305	2†	0.3†	
Si IV + O IV]	1,400	5	0.8	
N IV]	1,486	4†	0.6†	Possible
C IV	1,549	58†	7.8†	Possible small contribution from Si II $\lambda 1526$
[Ne IV]	1,602	6†	0.7†	Possible
He II	1,640	4†	0.5†	Possible
C III]	1,909	10	1.2	
Mg II	2,800	17	1.6	

†Uncertain.

Table 6 Absorption lines in the spectrum of 3C273

Identification	λ_{obs} (Å)	Comments
O I	1,302	Galactic
C II	1,335	Galactic
?	1,374†	Possible
Al II	1,671	Galactic, possible

†Uncertain

to be responsible for the strong optical Fe II emission. Detailed model predictions are required to resolve this question. (See note added in proof.)

The very efficient conversion of the UV continuum into optical emission required if resonance fluorescence is the principle excitation mechanism implies 0.5–1.0 of the sky as seen from the continuum source is covered by optically-thick clouds⁸². This conclusion also seems consistent with the large equivalent width of H β generally observed^{40,82}. However, we would then expect $W_{\lambda_0}(\text{Ly}\alpha) \approx 700 \text{ \AA}$ for a continuum slope $\alpha = 1^{40}$, but the observed values, as for 3C273, are about an order of magnitude smaller. The known general absence of Lyman continuum absorption in the rest frame of high redshift QSOs⁸⁶ implies that <0.1 of the continuum radiation is intercepted by optically thick clouds. This is consistent with the observed Ly α equivalent widths (if Ly α has not been strongly attenuated) and also with model calculations embodying other UV lines observed in high redshift QSOs⁵², yet is inconsistent with the H β (and Fe II) strengths which become anomalously high for a covering factor <0.1 .

Narrow absorption lines resembling the interstellar lines of our Galaxy are frequently observed in the spectra of QSOs and BL Lac objects. In general, these lines may be grouped into one or more systems corresponding with the apparent wavelengths of transitions in a physically plausible group of atoms or ions exhibiting the same, although arbitrary, redshift^{88,89}. Such absorption systems range in redshift from being comparable with the associated emission line redshift, to having very much smaller values which in some cases give a velocity relative to the QSO exceeding 0.6c. The general character of the absorption lines does not seem to differ for systems widely ranging in redshift, and velocity relative to the QSO. However, for QSOs of sufficient redshift to bring the Ly α emission line into the observable optical range, it is striking that most of the absorption lines fall shortwards of Ly α emission, and that the bulk of these cannot be grouped into convincing absorption systems. It is generally accepted that these unidentified lines represent Ly α absorption in gas extending to lower redshifts than the QSO against which they are observed and for which lines of other elements are too weak to be detected⁹⁰. In support of this is the frequent identification in these cases of one or more other members of the Lyman series^{91,92}.

The question of the origin of QSO absorption lines has been greatly discussed^{7,93–106}. In principle, the narrow-line absorbing regions may be intrinsic to the QSOs, having been somehow ejected at velocities up to the very high values observed, or they may be cosmologically distributed in direct association with galaxies or as intergalactic clouds. The outflow hypothesis has several difficulties, principally the implication of implausibly large mass and momentum flow relative to the QSO^{97–100}, and the balance of evidence now strongly favours the cosmological hypothesis. On the other hand, the very broad-line absorption present in some QSOs is probably intrinsic⁷.

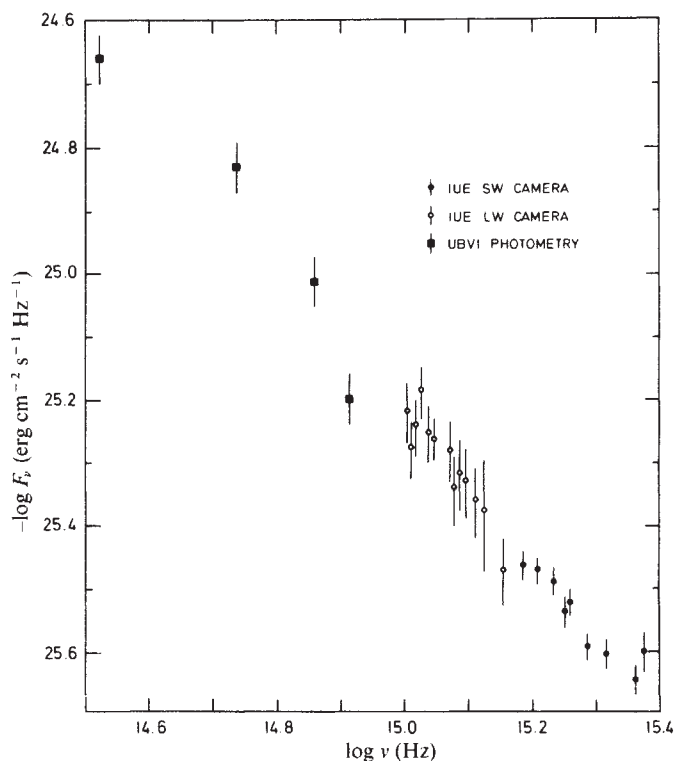
From the observed frequency of absorption systems containing heavy elements, if related to galaxies, the effective size of a galaxy seen in absorption must be several times its arbitrary Holmberg dimensions^{7,94,106,107}. However, as pointed out by Bahcall and Spitzer¹⁰⁸ the gas density required to produce absorption lines is so low that it would not be surprising if the maximum dimensions for measurable absorption in a line were much greater than the dimensions detected in other ways, and they suggested that the observed lines could be produced in

extended low density haloes of normal galaxies. Direct evidence that heavy element absorption lines in QSO spectra can be caused by the interception of the outer regions of spiral galaxies well outside the Holmberg dimensions comes from the detection by Boksenberg and Sargent¹⁰⁹ of Ca II absorption lines in the spectrum of the QSO 3C232 ($z_{\text{em}} = 0.5303$) at the same redshift as the galaxy NGC3067 which lies 1.9 arc min away in the plane of the sky, or about 1.5 Holmberg radii (using $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$), following the discovery by Haschick and Burke¹¹⁰ of H21 cm absorption in this system. Taking into account the inclination of NGC3067, the line of sight intersects the plane of the galaxy at a radial distance greater than 6 Holmberg radii out from the centre. However, the characteristics of the absorption system point to it arising in the outer halo of NGC3067 not the disk extension¹⁰⁹.

For 3C273 we find absorption lines only at zero redshift (Table 6), presumably arising in our Galaxy. Corresponding galactic Ca II absorption has been reported by Greenstein¹¹¹. Wagoner¹¹² has computed the probability that light received from a distant QSO be intercepted by an intervening galaxy and show absorption lines. Assuming an effective cross-section 10 times the area defined by the Holmberg dimensions, Wagoner's data still give only 0.05 as the probability of interception along the line of sight to 3C273, which is consistent with our null detection. This further supports the cosmological hypothesis for the absorbing material. On the outflow hypothesis the probability of detecting absorption systems is not expected to depend strongly on the redshift of the QSO.

Turning to the consideration of the Ly α absorption lines normally seen optically in the spectra of high redshift QSOs, even if the resolution of our 3C273 spectrum were not sufficient to separate these lines, if present, we would expect them to blend into a marked depression of the apparent continuum shortwards of Ly α emission. Neither we nor Davidsen *et al.*⁴² observe such a depression. The observed number density of the Ly α absorption lines seems to increase quite rapidly with redshift, with objects of $z_{\text{em}} \sim 2$ showing a markedly lower absorption line density than those of $z_{\text{em}} \sim 3$. Our null detection for 3C273 follows this trend and again supports the cosmological hypothesis. There is

Fig. 8 B2 1101+38 continuum fluxes. The IUE data are averaged over 50 Å bands. The UBVI photometry are approximately concurrent with the IUE measurements (see Table 1). Errors are estimated only from the 1σ statistical uncertainty in signal. No reddening correction is applied.



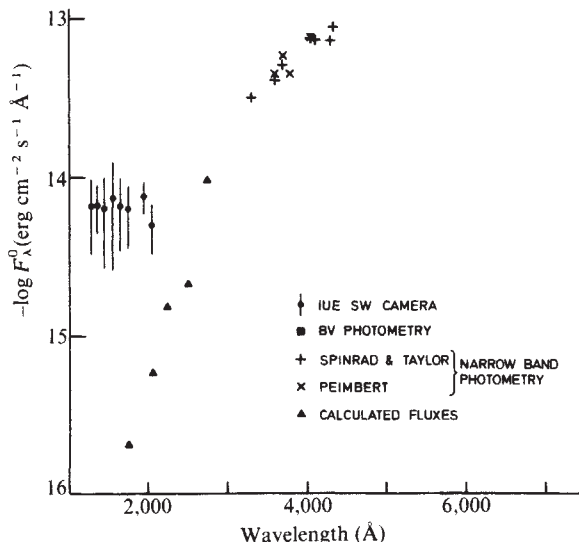


Fig. 9 Spectrophotometry and photometry of the central region of the spiral galaxy M81. The IUE short wavelength measurements are averaged over 100 Å bands. The errors define 1σ statistical uncertainties in signal only. The BV photometry (from ref. 127) and the narrow band photometry are reduced to the IUE observing aperture. The data are de-reddened using $E(B - V) = 0.06$. The derivation of the calculated fluxes is described in the text.

evidence that the Lyα absorption systems represent in part very extended outer regions of galaxies where the density is too low to allow star formation and the gas still has primordial composition, and in part truly intergalactic gas not bound to galaxies⁷.

B2 1101 + 38

B2 1101 + 38 (= Markarian 421) is a radio-faint galaxy¹¹⁷ identified as a BL Lac object^{118,119} and also is an X-ray source^{120,144}. It is among a few BL Lac objects observed to be surrounded by a nebulosity. Three absorption features in the spectrum of the nebulosity are consistent with being Ca II K, H and the G-band at $z = 0.03$ (refs 118, 121) and may be interpreted as originating from the main body of an elliptical galaxy having an extremely bright nucleus. BL Lacertae itself is similar, with $z = 0.07$ (refs 122–125). Five galaxies within 15' of B2 1101 + 38 are found also to have $z = 0.03$, showing that this object belongs to a cluster of galaxies¹²¹.

BL Lacertae objects in general appear similar to variable QSOs except that they lack strong optical emission lines. For a review of BL Lac objects see Stein *et al.*¹²⁶. Attempts to explain the absence of emission lines are usually based on (1) the supposition of a lack of UV photons to ionise a surrounding gas; (2) the absence of any such gas; or (3) a combination of both.

In our extension into the far UV there are still no emission-lines evident in the spectrum of B2 1101 + 38. There is, however, a strong non-thermal continuum. In Fig. 8 we show the IUE continuum measurements and include our UBVI photometry (see Table 1). Assuming a power law spectrum of the usual form $F_\nu \propto \nu^{-\alpha}$ we obtain $\alpha = 1.2$ as a mean value over the observed range. The radio spectrum is flatter, with $\alpha = 0.26$. As there is no sign of Lyα absorption at $z = 0.03$ (which would be clear of the geocoronal Lyα emission) we may assume that the observed UV spectrum extends well into the Lyman continuum. This is supported by the measurements of the possible X-ray counterpart by Ricketts *et al.*¹²⁰ which indicate a steady component of 0.7 Ariel 5 count s⁻¹ (sometimes brightening by up to a factor 20) while extrapolation of our continuum predicts 0.5 counts s⁻¹. There may be slight reddening, as our data are too noisy to confirm the absence of a weak absorption feature at local or redshifted λ2200, but there is no strong reddening so indicated.

If photoionisation is the source of line emission we may calculate the equivalent width of Hβ in terms of the photoionising continuum, assuming that the gas which gives rise to line emission is optically thick and completely surrounds the

continuum source. Using the expression given by Penston and Fosbury⁸⁴ (taking a continuum high-frequency cutoff at λ228) we obtain $W_\lambda(H\beta) \approx 50 \text{ Å}$ for $\alpha = 1.20$. (There is an omission in their expression as given; the correct expression is

$$W_\lambda(H\beta) \approx 564 \left(9.64 \frac{0.553}{K} \right)^{-\alpha} \left[\frac{K^{-\alpha} - 1}{\alpha} \right]$$

where $K = 0.553$ or 0.250 with the continuum high-frequency cutoff at λ504 or λ228 respectively, and α is defined above). If the intensity ratio of Lyα to Hβ is ≈ 3 , as is now commonly observed or deduced (see above) and using

$$I(\text{Ly}\alpha)/I(\text{H}\beta) = (4861/1216)^{2-\alpha} W_\lambda(\text{Ly}\alpha)/W_\lambda(\text{H}\beta),$$

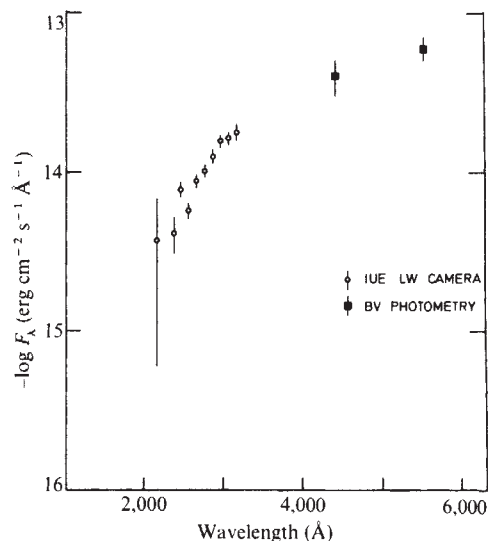
then $W_\lambda(\text{Ly}\alpha) \approx 50 \text{ Å}$. Our observed upper limit is $W_\lambda(\text{Ly}\alpha) \leq 10 \text{ Å}$. In the previous section we have discussed the conflicting evidence regarding complete (or nearly complete) coverage of the photoionising source. However, the Hβ strength normally observed in QSOs and Seyfert galaxies is well explained in this way. Thus we may conclude at least for B2 1101 + 38, and probably for BL Lac objects in general, that the absence of emission lines is due not to a deficiency of ionising flux, but of gas in the physical or geometrical state normally found in the line-emitting regions of low redshift QSOs. A similar conclusion follows from the lack of detectable X-ray absorption in such objects¹⁴⁴.

From Baldwin's⁴³ empirical relationship coupling $W_{\lambda_0}(\text{C IV})$ with $L_\nu(1450)$, discussed above for NGC4151, NGC1068 and 3C273, we expect from our continuum flux that $W_{\lambda_0}(\text{C IV}) = 590 \text{ Å}$, but our observed upper limit is $W_{\lambda_0}(\text{C IV}) \leq 15 \text{ Å}$. Again we have a serious departure from Baldwin's empirical relation; it is interesting that B2 1101 + 38 has a flat radio spectrum and Baldwin finds his correlation is particularly tight for QSOs with such spectra.

M81

The central region of the well-studied spiral galaxy M81 was observed only with the short wavelength spectrograph on IUE (Table 1). The signal obtained was low, and to increase the signal-to-noise ratio we binned the data into 100 Å bands, as shown in Fig. 9 which also shows the BV photometry of Brandt *et al.*¹²⁷ and the narrow band photometry of Spinrad and Taylor¹²⁸ and Peimbert¹²⁹. The Spinrad and Taylor results were normalised to Peimbert's at λ5360. These are all reduced effectively to the same observing aperture as used for the IUE data. All the observed data are de-reddened using $E(B - V) = 0.06$ (ref. 128), and the extinction law of Code *et al.*³⁸ in the UV.

Fig. 10 Spectrophotometry and photometry of the central region of the giant elliptical galaxy M87. The long wavelength IUE data are averaged over 100 Å bands and include only 1σ statistical uncertainties in signal. The photometry are from refs 137 and 139 and are reduced to the IUE observing aperture. No reddening correction is applied.



We also show a theoretical extension of the optical data into the UV calculated from the stellar population of Spinrad and Taylor¹²⁸ and the appropriate UV stellar fluxes from the TD1 S2/68 instrument¹⁰.

The observed IUE spectrum clearly does not arise from the assumed stellar population, whose emission is dominantly from G-type subgiants, but in magnitude agrees well with Peimbert's¹²⁹ estimate of the fluxes required from a hot star population (late O-type) needed to explain the intensity of the observed H α and [O III] emission in this object. However, our flux distribution has the wrong slope for Peimbert's assumed stellar population, and is fitted better by a mean A0 spectrum, but this will not be able to provide the required photoionising flux. Alternatively, the observations are well-explained by a non-thermal continuum source of spectral slope $\alpha \approx 2$. The radio nucleus of M81 is known to be variable and is about 10^4 times more luminous than the compact radio source at the galactic centre⁸⁰.

M87

The peculiar EO galaxy M87 (=NGC4486=Arp 152=3C274=Virgo A) is an active (core-halo) radio source with compact central components^{130,131}. It also has compact and extended X-ray emission¹³⁴⁻¹³⁶, and there is evidence for ejection of mass from the nucleus in the form of optical jets^{132,133}, and for a massive black hole in the nucleus^{137,138}.

M87 was observed by IUE only with the long wavelength spectrograph. Again we binned the data into 100Å bands to increase the signal-to-noise ratio. The results are shown in Fig. 10. We include BV photometry from de Vaucouleurs and Nieto¹³⁹ and Young *et al.*¹³⁷, reduced to the same observing aperture as for the IUE data.

From our UV data, an early G-type spectrum is indicated for the nuclear region of M87. Late G-type to early K-type (dominant in the visible flux) is too cool, as for this an absorption edge cuts off the flux below $\sim \lambda 2640$. F-type subgiants can be ruled out as the slope is wrong and the absorption dip at $\lambda 2400$ present in F-type stars¹⁰ is not evident in the IUE spectrum. Actually, there is marginal evidence for an emission peak (perhaps [Ne IV] $\lambda 2424$). There may also be some contribution from a non-thermal continuum, which is expected for this object.

We thank J. McClintock for optical spectra of NGC4151 and 3C273, P. R. Wesselius and D. J. Carnochan for results before publication, and D. C. Morton and J. B. Oke for helpful comments.

Note added in proof: Recent computations^{145,146} suggest observational evidence does point strongly to collisional excitation as the dominant mechanism producing Fe II optical emission lines.

1. Boggess, A. *et al.* *Nature* **275**, 361-369 (1978).
2. Bless, R. C., Code, A. D. & Fairchild, E. T. *Astrophys. J.* **203**, 410 (1976).
3. Beckmann, F. *Astr. Astrophys.* **60**, 1 (1977).
4. Strongylis, G. J. & Bohlin, R. C. *Astrophys. J.* (in the press).
5. Heap, S. R. *et al.* *Nature* **275**, 369-376 (1978).
6. Code, A. D., Holm, A. V. & Bottemiller, R. L. *Astrophys. J.* (in the press).
7. Boksenberg, A. *Physica Scripta* **17**, 205 (1978).
8. Wu, C.-C. *Astrophys. J. Lett.* **217**, L117 (1977).
9. Wesselius, P. R. & Koester, D. *Astr. Astrophys.* (in the press).
10. Jamar, C. *et al.* *ESA SR27*, Ultraviolet Bright-Star Spectrophotometric Catalogue (1976).
11. Weedman, D. W. *A. Rev. Astr. Astrophys.* **15**, 69 (1977).
12. Kachikyan, E. Ye. E. & Weedman, D. W. *Astrofizika* **7**, 389 (1973).
13. Ives, J. C., Sanford, P. W. & Penston, M. V. *Astrophys. J. Lett.* **207**, L159 (1976).
14. Wu, C.-C. & Weedman, D. W. *Astrophys. J.* **223**, 798 (1978).
15. Fitch, W. S., Pacholczyk, A. G. & Weymann, R. J. *Astrophys. J. Lett.* **150**, L67 (1967).
16. Penston, M. V., Penston, M. J., Selmes, R. A., Becklin, E. E. & Neugebauer, G. *Mon. Not. R. Astr. Soc.* **169**, 357 (1974).
17. Rieke, G. H. & Low, F. J. *Astrophys. J. Lett.* **176**, L95 (1972).
18. Di Cocco, G., *et al.* *Nature* **270**, 319 (1977).
19. Heesch, D. S. & Wade, C. M. *Astr. J.* **69**, 277 (1964).
20. de Bruyn, A. G. & Wilson, A. S. *Astr. Astrophys.* **64**, 433 (1978).
21. Oke, J. B. & Sargent, W. L. W. *Astrophys. J.* **151**, 807 (1968).
22. Anderson, K. *Astrophys. J.* **162**, 743 (1970).
23. Netzer, H. *Mon. Not. R. Astr. Soc.* **169**, 579 (1974).
24. Boksenberg, A. *et al.* *Mon. Not. R. Astr. Soc.* **173**, 381 (1975).
25. Boksenberg, A. & Penston, M. V. *Mon. Not. R. Astr. Soc.* **177**, 127P (1976).
26. Searle, L. & Sargent, W. L. W. *Astrophys. J.* **153**, 1003 (1968).
27. Osterbrock, D. E. *Proc. Natn. Acad. Sci. U.S.A.* **75**, 540 (1978).
28. Osterbrock, D. E., Koski, A. T. & Phillips, M. M. *Astrophys. J.* **206**, 898 (1976).
29. Ptak, R. L. & Stoner, R. E. *Astrophys. J.* **185**, 121 (1973).
30. Ulrich, M.-H. *Astrophys. J. Lett.* **171**, L37 (1972).
31. Blumenthal, G. R. & Mathews, W. G. *Astrophys. J.* **198**, 517 (1975).
32. Anderson, K. S. *Astrophys. J.* **189**, 195 (1974).
33. Weymann, R. J. & Cromwell, R. *Proc. IAU Sym. No. 44, External Galaxies and Quasi-Stellar Objects* (ed. Evans, D. E.) 155 (Reidel, Dordrecht 1972).
34. Stein, W. A. & Weedman, D. W. *Astrophys. J.* **205**, 44 (1976).
35. Wampler, E. J. *Astrophys. J.* **164**, 1 (1971).
36. Johnson, H. L. & Borgman, J. *B.A.N.* **17**, 115 (1963).
37. Nandy, K., Thompson, G. I., Jamar, C., Monfils, A. & Wilson, R. *Astr. Astrophys.* **44**, 195 (1975).
38. Code, A. D., Davis, J., Bless, R. C. & Hanbury Brown, R. *Astrophys. J.* **203**, 417 (1976).
39. Netzer, H. *Mon. Not. R. Astr. Soc.* **171**, 395 (1975).
40. Baldwin, J. A. *Mon. Not. R. Astr. Soc.* **178**, 67P (1977).
41. Hyland, A. R., Becklin, E. E. & Neugebauer, G. *Astrophys. J. Lett.* **220**, L73 (1978).
42. Davidsen, A. F., Hartig, G. F. & Fastie, W. G. *Nature* **269**, 203 (1977).
43. Baldwin, J. A. *Astrophys. J.* **214**, 679 (1977).
44. Osmer, P. S. *Astrophys. J.* **214**, 1 (1977).
45. Grandi, S. A. *Astrophys. J.* **221**, 501 (1978).
46. Davidsen, A. F. & Hartig, G. F. preprint (1978).
47. Viotti, R. *Mon. Not. R. Astr. Soc.* **177**, 617 (1976).
48. Baldwin, J. A., Burke, W. L., Gaskell, C. M. & Wampler, E. J. *Nature* **273**, 431 (1978).
49. Tananbaum, H. *et al.* *Astrophys. J.* **223**, 74 (1978).
50. Zirin, H. *Astrophys. J. Lett.* **222**, L105 (1978).
51. Shields, G. A. *Astrophys. J.* **204**, 330 (1976).
52. Baldwin, J. A. & Netzer, H. preprint.
53. Davidson, K. *Astrophys. J.* **218**, 20 (1977).
54. Osterbrock, D. E. & Parker, R. A. *R. Astr. Soc. Astrophys. J.* **141**, 892 (1965).
55. Wampler, E. J. *Astrophys. J.* **164**, 1 (1971).
56. Shields, G. A., Oke, J. B. & Sargent, W. L. W. *Astrophys. J.* **176**, 75 (1972).
57. Shields, G. A. & Oke, J. B. *Astrophys. J.* **197**, 5 (1975).
58. Neugebauer, G. *Physica Scripta* **17**, 149 (1978).
59. Walker, M. F. *Astrophys. J.* **151**, 71 (1968).
60. Adgie, R. L. *Nature* **204**, 1028 (1964).
61. Cohen, M. H., Gundermann, E. J. & Harris, D. E. *Astrophys. J.* **150**, 767 (1967).
62. Thompson, R. I., Lebofsky, M. J. & Rieke, G. H. *Astrophys. J. Lett.* **222**, L49 (1978).
63. Rieke, G. H. & Low, F. J. *Astrophys. J. Lett.* **177**, L115 (1972).
64. Kleinmann, D. E., Gillett, F. C. & Wright, E. L. *Astrophys. J.* **208**, 42 (1976).
65. Boksenberg, A. & Sargent, W. L. W. (in preparation).
66. Collin-Souffrin, S. *Physica Scripta* **17**, 293 (1978).
67. Koski, A. T. *Astrophys. J.* **223**, 56 (1978).
68. Angel, J. R. P., Stockman, H. S., Woolf, N. J., Beaver, E. A. & Martin, P. G. *Astrophys. J.* **206**, L5 (1976).
69. Strittmatter, P. A. *Physica Scripta* **17**, 145 (1978).
70. Greenstein, J. L. & Schmidt, M. *Astrophys. J.* **140**, 1 (1964).
71. Oke, J. B. *Astrophys. J.* **141**, 6 (1965).
72. Wampler, E. J. & Oke, J. B. *Astrophys. J.* **148**, 695 (1965).
73. Boksenberg, A., Shortridge, K., Fosbury, R. A. E., Penston, M. V. & Savage, A. *Mon. Not. R. Astr. Soc.* **172**, 289 (1975).
74. Baldwin, J. A. *Astrophys. J. Lett.* **196**, 191 (1975).
75. Baldwin, J. A. *Astrophys. J. Lett.* **201**, 26 (1975).
76. Bowyer, C. S., Lampton, M., Mack, J. & de Mendonca, F. *Astrophys. J. Lett.* **161**, L1 (1970).
77. Kellogg, E. *et al.* *Astrophys. J. Lett.* **165**, L49 (1971).
78. Rieke, G. H. & Low, F. J. *Astrophys. J. Lett.* **176**, L95 (1972).
79. Giacconi, R. *Physica Scripta*, **17**, 159 (1978).
80. Kellerman, K. I. *Physica Scripta*, **17**, 257 (1978).
81. Phillips, M. M. *Lick Observatory Bull.* No. 802 (1978).
82. Phillips, M. M. *Lick Observatory Bull.* No. 808 (1978).
83. Grasdalen, G. L. *Astrophys. J.* **208**, L11 (1976).
84. Penston, M. V. & Fosbury, R. A. E. *Mon. Not. R. Astr. Soc.* **183**, 479 (1978).
85. Burbidge, E. M. *Physica Scripta* **17**, 165 (1978).
86. Osmer, P. S. & Smith, M. G. *Astrophys. J.* **213**, 607 (1977).
87. Chan, Y.-W. T. & Burbidge, E. M. *Astrophys. J.* **198**, 45 (1975).
88. Bahcall, J. N. *Astrophys. J.* **153**, 679 (1968).
89. Aaronson, M., McKee, C. F. & Weisheit, J. C. *Astrophys. J.* **198**, 13 (1975).
90. Lynds, C. R. *Astrophys. J. Lett.* **164**, L73 (1971).
91. Baldwin, J. A. *et al.* *Astrophys. J.* **193**, 513 (1974).
92. Young, P. J., Sargent, W. L. W., Boksenberg, A., Carswell, R. F. & Whelan, J. A. J. preprint.
93. Bahcall, J. N. *Astr. J.* **76**, 283 (1971).
94. Bahcall, J. N. *Astrophys. J. Lett.* **200**, L1 (1975).
95. Bahcall, J. N. *Physica Scripta* **17**, 229 (1978).
96. Burbidge, E. M. & Burbidge, G. R. *Astrophys. J.* **202**, 287 (1975).
97. Goldreich, P. & Sargent, W. L. W. *Comments Astr. Astrophys.* **6**, 133 (1976).
98. Williams, R. E. & Weymann, R. J. *Astrophys. J. Lett.* **207**, L143 (1976).
99. Wolfe, A. M., Broderick, J. J., Condon, J. J. & Johnston, K. J. *Astrophys. J. Lett.* **208**, L47 (1976).
100. Wolfe, A. M., Broderick, J. J., Condon, J. J. & Johnston, K. J. *Astrophys. J.* **222**, 752 (1978).
101. Weymann, R. J. & Williams, R. E. *Physica Scripta* **17**, 217 (1978).
102. Goldreich, P. *Physica Scripta* **17**, 201 (1978).
103. Burbidge, E. M. *Physica Scripta* **17**, 201 (1978).
104. Burbidge, G. R. *Physica Scripta* **17**, 237 (1978).
105. Burbidge, E. M., Smith, H. E., Weymann, R. J. & Williams, R. E. *Astrophys. J.* **218**, 1 (1977).
106. Burbidge, G. R., O'Dell, S. L., Roberts, D. H. & Smith, H. E. *Astrophys. J.* **218**, 33 (1977).
107. Roberts, D. H., Burbidge, E. M., Burbidge, G. R., Crowne, A. H., Junkkarinen, V. T. & Smith, H. E. preprint.
108. Bahcall, J. N. & Spitzer, L. *Astrophys. J. Lett.* **156**, L63 (1969).
109. Boksenberg, A. & Sargent, W. L. W. *Astrophys. J.* **220**, 42 (1978).
110. Haschick, A. D. & Burke, B. F. *Astrophys. J. Lett.* **200**, L137 (1975).
111. Greenstein, J. L. *Astrophys. J.* **152**, 431 (1968).
112. Wagoner, R. V. *Astrophys. J.* **149**, 465 (1967).
113. Wampler, E. J. & Oke, J. B. *Astrophys. J.* **148**, 695 (1967).
114. Shields, G. A. *Astrophys. J.* **191**, 309 (1974).
115. Boksenberg, A. & Netzer, H. *Astrophys. J.* **212**, 37 (1977).
116. Adams, T. F. *Astrophys. J.* **196**, 675 (1975).
117. Colla, G. *et al.* *Astr. Astrophys. Suppl.* **20**, 1 (1975).
118. Ulrich, M.-H., Kinman, T. D., Lynds, C. R., Rieke, G. H. & Ekers, R. D. *Astrophys. J.* **198**, 261 (1975).
119. Ulrich, M.-H. *Astrophys. Lett.* **14**, 89 (1973).
120. Ricketts, M. J., Cooke, B. A. & Pounds, K. A. *Nature* **259**, 546 (1976).
121. Ulrich, M.-H. *Astrophys. J. Lett.* **222**, L3 (1978).
122. Oke, J. B. & Gunn, J. E. *Astrophys. J. Lett.* **189**, L5 (1974).
123. Thuan, T. X., Oke, J. B. & Gunn, J. E. *Astrophys. J.* **201**, 45 (1975).

124. Miller, J. S. & Hawley, S. A. *Astrophys. J. Lett.* **212**, L47 (1977).
 125. Miller, J. S., French, H. B. & Hawley, S. A. *Astrophys. J. Lett.* **219**, L85 (1978).
 126. Stein, W. A., O'Dell, S. L. & Strittmatter, P. A. *A. Rev. Astr. Astrophys.* **14**, 173 (1976).
 127. Brandt, J. C., Kalinowski, J. K. & Roosen, R. J. *Astrophys. J. Suppl.* **24**, 421 (1972).
 128. Spinrad, H. & Taylor, B. J. *Astrophys. J. Suppl.* **22**, 445 (1971).
 129. Peimbert, M. *Astrophys. J.* **154**, 33, (1968).
 130. Cohen, M. H. *et al. Astrophys. J. Lett.* **158**, L83 (1969).
 131. Kellerman, K. I. *et al. Astrophys. J. Lett.* **179**, L141 (1973).
 132. Felten, J. E. *Astrophys. J.* **151**, 861 (1968).
 133. Felten, J. E., Arp, H. C. & Lynds, C. R. *Astrophys. J.* **159**, 415 (1970).
 134. Gorenstein, P., Topka, K., Fabricant, D. & Harnden, F. R. *Bull. Am. Ass. Soc.* **8**, 553 (1976).
 135. Catura, R. C. & Acton, L. W. *Bull. Am. Astr. Soc.* **8**, 554 (1976).
 136. Malina, R., Lampton, M. & Bowyer, S. *Astrophys. J.* **209**, 678 (1976).
 137. Young, P. J., Westphal, J. A., Kristian, J., Wilson, C. P. & Landauer, F. P. *Astrophys. J.* **221**, 721 (1978).
 138. Sargent, W. L. W. *et al. Astrophys. J.* **221**, 731 (1978).
 139. de Vaucouleurs, G. & Nieto, J.-L. *Astrophys. J.* **220**, 449 (1978).
 140. Anderson, K. S. & Kraft, R. P. *Astrophys. J.* **158**, 859 (1969).
 141. Cameron, A. in *Explosive Nucleosynthesis* (eds. Schramm, D. N., & Arnett, W. D. University of Texas Press, Austin, 1973).
 142. Osterbrock, D. E. *Astrophys. J.* **215**, 733 (1977).
 143. Jordan, C. in *Progress in Atomic Spectroscopy* (eds. Hanle and Kleinpoppen) (Plenum Oxford, in the press).
 144. Mushotzky, R. F. *et al.* preprint.
 145. Collin-Souffrin, S., Joly, M., Heidmann, N. & Dumont, S. *Astr. Astrophys.* (in the press).
 146. Collin-Souffrin, S., Dumont, S., Heidmann, N. & Joly, N. *Astr. Astrophys.* (in the press).

IUE observations of Solar System objects

A. L. Lane & E. Hamrick

Jet Propulsion Laboratory, Pasadena, California 91103

A. Boggess, D. C. Evans & T. R. Gull

Goddard Space Flight Center, Greenbelt, Maryland 20771

F. H. Schiffer III, B. Turnrose, P. Perry & A. Holm

Computer Sciences Corporation, Silver Spring, Maryland 20900

F. Macchetto

Astronomy Division, ESTEC, Noordwijk, Holland

P. M. Gondhalekar

SRC Appleton Laboratory, Slough, Berks, UK

G. E. Hunt & R. Wilson

University College London, Gower St, London WC1, UK

T. C. Owen

State University of New York at Stony Brook, New York 11790

H. W. Moos

Johns Hopkins University, Baltimore, Maryland 21218

M. G. Tomasko & T. Gehrels

University of Arizona, Tucson, Arizona 85721

R. Conway & C. A. Barth

University of Colorado, Boulder, Colorado 80302

During the scientific commissioning phase of IUE several spectra were acquired from objects residing in the Solar System. The activities focused on testing numerous parameters which would indicate the usefulness of IUE for planetary science. It seems that IUE can successfully tackle many important questions and the data analysis and interpretation of the initial set of observations has begun.

IUE PROVIDES an excellent way for acquiring UV spectra in the wavelength region shortwards of the Earth's ozone cutoff (about 3,100 Å). The UV vidicons, when properly erased and prepared for use, can be exposed for several hours on very weak planetary emissions. This integration time, and the ability of the spacecraft to be adjusted to track on planets or their natural satellites for extended periods, combine to provide an overall capability which is well beyond the reach of sounding rockets. The low resolution channel (linear spectrum) has a spectral resolution of 7–8 Å FWHM, much like the classical rocket and satellite spectrometers now in use. The high resolution mode (echelle spectrum) gives a spectral resolution in the range 0.1–0.2 Å FWHM. (The scientific instrument is described by Boggess *et al.*¹) Either the large 10 × 20 arc s slot or the 3 arc s circular aperture can be used for studies requiring spatial resolution, if the tracking can be adjusted and maintained to ~1 arc s during an exposure. The sample spectra which were recorded during the IUE commissioning phase show some of the possibilities for future studies with Solar System objects. Table 1 represents the wide range of objects examined. The efforts centred around acquiring exposure times for reasonable signal-to-noise ratio spectra and becoming familiar with the IUE capability. No definitive study of any one object was performed, but an overall analysis of what further studies would be possible was accomplished.

Figure 1 reproduces a single long-wavelength spectrum of Mars, acquired through the small, 3 arc s aperture. Useable information is obtained down to 2,200 Å. The decrease in signal strength longwards of 2,900 Å is the result of the threshold of

the caesium telluride photocathode in the UV quantum converter which is in front of the vidicon faceplate. The spectrum was obtained by placing the small aperture over the 8 arc s diameter disk of Mars, and using the fine error sensor (FES) to lock onto the 'spill-over' light. The exposure was short enough (15 s) to not require trimming of the spacecraft gyros to the non-sidereal motion of Mars. The strong Mg I and Mg II solar absorption lines are easily detected at 2,852 and 2,795 and 2,802 Å, respectively. Just longwards of 2,900 Å, the spectrum saturated slightly, but the data became valid again at 2,960 Å. Although the signal is not strong in the 2,400–2,600 Å region, no pronounced absorption band is seen around 2,550 Å where the Hartley band system of ozone has its maximum absorption. If substantial ozone were present in the martian atmosphere, one would expect to see an absorption bowl in the Mars spectrum². To enhance the detectability, within the limits of the statistics of a single spectrum, the Mars spectrum was compared with a lunar spectrum also obtained with IUE. This is displayed in Fig. 2. No apparent minimum is observed and the slope of the ratio represents the albedo differences of the two bodies. The Mars spectrum was recorded while its northern hemisphere was in late spring ($L_s = 70^\circ$) and the north polar region was tipped towards Earth, substantially obscuring the south polar region. In

Table 1 Objects from the Solar System observed using IUE

Object	Wavelength range	Dispersion
Mars	Long	High and low
	Short	Low
Jupiter	Long	High and low
	Short	Low
Saturn	Long	Low
Uranus	Long	Low
	Short	Low
Io	Long	Low
Ganymede	Long	Low
Titan	Long	Low
Moon	Long	High and low
	Short	Low

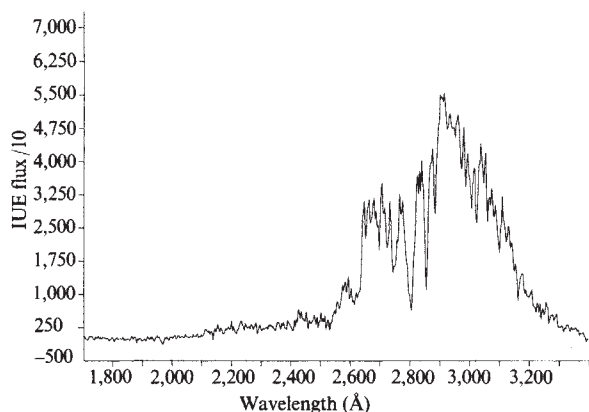


Fig. 1 Long wavelength, low resolution spectrum of Mars recorded on April 1978 with a small aperture. The system response function has not been removed. Exposure time is 15 s.

this configuration the northern hemisphere contributed ~60–65% of the total reflected light. The water vapour content of the Mars lower atmosphere becomes considerable during the late spring and its subsequent photodissociation leads to a process which inhibits ozone formation^{3–7}. Mariner 9 observed about 10 μ atm of ozone near the north pole of Mars at the equivalent seasonal time in 1972 (ref. 6), but detected no ozone near the equatorial region. Because the aperture was locked onto the centroid of light at least $1\frac{1}{2}$ arc s of the north polar region of Mars was excluded from the spectrometer. It is highly unlikely that ozone would have extended southwards from the north polar region to latitudes as far south as $+45^\circ$. Thus, if the seasonal characteristics in 1978 were similar to those in 1972, virtually no ozone should have been in the martian atmosphere which was included in the aperture. Confirming studies of the ozone seasonal characteristics will have to wait until about December 1978 when the winter–spring transition can be observed in the southern hemisphere and when one should expect to detect atmospheric ozone.

Acquiring data from the jovian system requires special considerations and techniques. Figure 3 displays the reflected light spectrum of Io (J1), Jupiter's innermost galilean satellite. The region near 2,900 Å is slightly saturated and fiducial marks on the camera faceplate account for the two apparent absorption features near 2,580 and 3,305 Å. An observation time was chosen when Io and Ganymede were simultaneously near elongation, as seen from Earth, and had minimal differential motions for several hours. The tracking procedure required FES acquisition of Jupiter, which is much too large and too bright to be useable for the tracker, followed by FES lock-up on Ganymede. The spacecraft gyros were trimmed to the Ganymede celestial motion and Io was subsequently manoeuvred into the large slot. The FES continued to lock on Ganymede while the 10 min exposure was in progress. Direct tracking on Io with the 'spill-over' light surrounding the large or small aperture (the

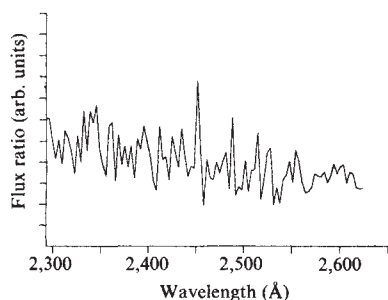


Fig. 2 Ratio of the spectrum of Mars (Fig. 1) to that of the Moon. Both spectra were recorded with IUE and the ratio was constructed from data which had received identical processing. The flux ratio is in arbitrary, linear units. No ozone absorption band is observed at 2,550 Å.

method used for Mars) is generally not possible because the scattered light from Jupiter is about equal to the 'spill-over' light from Io at Io's position in the tracker field. Europa (J2) may also have to be handled by this differential guiding technique, but Ganymede and Callisto are sufficiently removed from the visible scattered light domain of Jupiter to be directly tracked most of the time. Whenever any satellite is too close to Jupiter, within ~7 jovian radii, differential tracking will be necessary. This condition will also be true for satellite studies of the saturnian system. The spectrum of Io has not undergone extensive analysis because the data were acquired during a high radiation background portion of the IUE orbit and more background noise removal is necessary before proceeding further. Nonetheless, acquisition of Io's spectrum certainly demonstrates one of the capabilities of IUE.

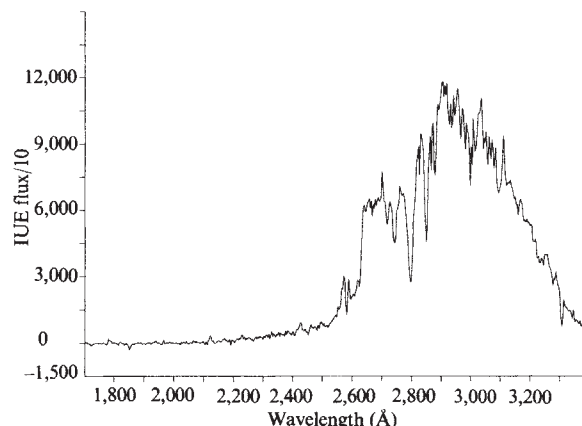


Fig. 3 Long wavelength, low resolution spectrum of Io using a large slot. The system response function has not been removed. Exposure time is 10 min.

The observation of Solar System objects with IUE is more complicated than that required for the classes of stars and galaxies for which IUE was originally designed. The significant differences are: (1) the range of visual magnitude brightness is very large, from -4 for Venus to fainter than $+11$ for some satellites of Saturn. This complicates acquisition and tracking. (2) Most of the bodies in the Solar System move at non-sidereal rates. Special tracking procedures need to be developed to acquire spectra of planetary satellites, regions of the bright disks of Jupiter, Saturn, Mars, Venus and the Moon. (3) All of the objects in this class echo the time variable solar spectrum, to first order. To examine the spectral range from 1,150 to 3,200 Å requires numerous exposures to cover a 10^5 brightness range with a detector system which is linear over about a ratio of 25 in object brightness in the UV. (4) Perhaps the most pressing complication facing IUE planetary observers is the acquisition of comparison solar spectra. Ideally they should be acquired by IUE itself so that all of the instrumental parameters are removed. IUE cannot look directly at the Sun, but the fully illuminated disk of the Moon near zero phase angle seems to be the best current choice for monitoring the solar spectrum. Some minor tracking difficulties still remain for accomplishing this, but they seem to be solvable. Alternate G type stars are being examined as another possibility to tie together solar spectral references.

All of these difficulties are temporary and excellent progress is being made to develop a powerful, useable system for the study of the Solar System in the UV. A brief inspection of the data acquired so far confirms that IUE is capable of providing measurements which can supply crucial answers to several of the more significant problems in planetary science.

1. Boggess, A. *et al.* *Nature* **275**, 356–361 (1978).
2. Lane, A. L., Barth, C. A., Hord, C. W. & Stewart, A. I. *Icarus* **18**, 102–108 (1973).
3. McElroy, M. B. & Hunten, D. M. *J. geophys. Res.* **75**, 1188 (1970).
4. McElroy, M. B. & McConnell, J. C. *J. atmos. Sci.* **28**, 879 (1971).
5. Parkinson, T. D. & Hunten, D. M. *J. atmos. Sci.* **29**, 1380 (1972).
6. Barth, C. A. *A. Rev. Earth planet. Sci.* **2**, 333 (1974).
7. Liu, S. C. & Donahue, T. M. *Icarus* **28**, 231 (1976).

articles

Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA

R. A. Laskey, B. M. Honda, A. D. Mills, & J. T. Finch

MRC Laboratory of Molecular Biology, Hills Road, Cambridge, UK

The nucleosome subunits of chromatin are assembled from histones and DNA by an acidic protein which binds histones. The nucleosome assembly protein has been identified and purified from eggs of Xenopus laevis.

DNA IN eukaryotic chromosomes is packed into a repeating chain of nucleosome subunits, each consisting of an octamer of histones complexed to approximately 200 base pairs of DNA (for reviews see refs 1 and 2). This article describes the identification and purification of a protein which binds histones and transfers them to DNA to form nucleosomes. The nucleosome assembly protein has been purified from unfertilised eggs of the frog *Xenopus laevis* by fractionating a cell-free nucleosome assembly system³. The cell-free system converts purified DNA into a regularly repeating chain of nucleosomes at a rate of 5×10^7 nucleosomes per s per egg³ using either the egg's endogenous histone pool⁴⁻⁶ or exogenous sources of purified histones. Assembly is not dependent on DNA synthesis³. The products obtained when DNA is incubated in the cell-free system were recognised as chains of nucleosomes by the criteria of: (1) sedimentation, (2) buoyant density, (3) beaded appearance in the electron microscope, (4) localised resistance to micrococcal nuclease, (5) histone dependence and (6) insertion of superhelical turns into relaxed circular DNA^{3,7}. We have used histone dependent insertion of superhelical turns into relaxed circular DNA as a rapid assay for nucleosome assembly because Germond *et al.*⁸ have shown that the number of superhelical turns is directly related to the number of nucleosomes on the DNA. This supercoiling activity should not be confused with the DNA gyrase of prokaryotes which also inserts superhelical turns into DNA⁹. In contrast to the activity we describe here, DNA gyrase requires ATP but does not require stoichiometric amounts of structural proteins such as histones, and in addition its activity is counteracted by nicking closing enzyme⁹.

An assembly factor is required for ordered interaction of histones with DNA

Nucleosomes can be reconstituted from separated DNA and histones by prolonged dialysis from $>1\text{M NaCl} \pm 5\text{M urea}$ (reviewed in ref. 2), but at physiological ionic strength mixtures of DNA and histones precipitate. Addition of small amounts of *Xenopus* egg homogenate prevents precipitation and results in nucleosome formation^{3,7}, suggesting that eggs contain an additional factor which is required for the ordered interaction of histones with DNA to form nucleosomes at physiological ionic strength.

The nucleosome assembly factor is a protein

To determine the nature of the assembly factor in eggs of *Xenopus laevis*, egg homogenates were incubated with a variety of specific degradative enzymes before assaying for assembly activity. The results are summarised in Table 1. The assembly factor is unlikely to contain a nucleic acid component since it is totally resistant to ribonuclease A, and micrococcal nuclease.

The assembly factor is also resistant to trypsin and chymotrypsin, but assembly is only observed if the proteases are inhibited by phenylmethylsulphonyl fluoride (PMSF) after digestion and if exogenous histones are added. (Exogenous nicking-closing enzyme is also required after digestion in order to visualise assembly by monitoring insertion of superhelical

Table 1 Sensitivity of the assembly factor to degradative enzymes

Enzyme preincubated with egg supernatant	Enzyme inhibitor	Assembly activity (using added histones)
Ribonuclease A ($500\text{ }\mu\text{g ml}^{-1}$)	Nil	+
Micrococcal nuclease ($1,500\text{ units ml}^{-1}$) + 2 mM CaCl_2	EDTA (5 mM)	+
Trypsin ($250\text{ }\mu\text{g ml}^{-1}$)	PMSF†	+
Chymotrypsin ($\leq 1\text{ mg ml}^{-1}$)	PMSF†	+
Proteinase K ($250\text{ }\mu\text{g ml}^{-1}$)	PMSF†	—
PMSF + proteinase K ($250\text{ }\mu\text{g ml}^{-1}$)	Inhibited before incubation	+

Xenopus egg homogenates were prepared and clarified as described previously³ except that $0.66\text{ volumes of }120\text{ mM KCl}$, 2 mM MgCl_2 , $20\text{ mM Tris-HCl pH }7.5$ were added after homogenisation and the $1,800\text{ g supernatant}$ was centrifuged twice at $145,000\text{ g (max)}$ for 30 min . $20\text{ }\mu\text{l}$ of clarified homogenate was incubated with each enzyme for 1 h at 25°C . Enzymes were then inhibited before addition of $1\text{ }\mu\text{g calf thymus histones}$, $0.5\text{ }\mu\text{g SV40 form I relaxed DNA}$ and nicking-closing enzyme which had been partially purified from eggs of *X. laevis* by chromatography on DEAE-cellulose and phosphocellulose columns¹². Assembly was assayed by incubation at 20°C for 3 h followed by agarose gel electrophoresis to reveal superhelical turns¹³. PMSF was dissolved in ethanol at 6 mg ml^{-1} and added to homogenates to give a final concentration of $150\text{ }\mu\text{g ml}^{-1}$.

* When micrococcal nuclease was not inhibited following pre-incubation, the SV40 DNA was completely degraded.

† When PMSF was omitted ^{35}S -labelled histones added to the incubation were digested to complete solubility in 5% trichloroacetic acid containing 0.25% sodium tungstate.

turns into relaxed DNA.) The resistance to trypsin and chymotrypsin suggests either that the assembly factor is not a protein or that it is a protein which lacks exposed basic and aromatic amino acids on its surface. To distinguish these two possibilities we tested its sensitivity to proteinase K (BDH, twice crystallised), a single protease which cleaves most types of accessible peptide bond and which is completely and irreversibly inhibited by PMSF¹⁰. The assembly activity is abolished by incubation with proteinase K. The destruction of activity is more rapid and more complete when the preincubation is performed above pH 8. Addition of histones with or without nicking-closing enzyme fails to restore any activity. When the proteolytic activity of proteinase K preparations is inactivated by PMSF before incubation with egg homogenates the assembly activity remains undiminished (Table 1), confirming that the loss of activity caused by proteinase K is due to proteolysis.

Identification and purification of a protein which co-purifies with assembly activity

To identify any proteins which co-purify with the assembly activity, clarified egg homogenates (see Table 1 legend) were initially fractionated by sedimentation through a 10–40% sucrose gradient. The distribution of polypeptides in the gradient fractions was examined by polyacrylamide gel electrophoresis of aliquots (Fig. 1a). Simultaneously each fraction was assayed for its ability to insert superhelical turns into SV40 circular DNA which had been relaxed by nicking-closing enzyme as described previously³. The number of superhelical turns is related to the number of nucleosomes on the DNA⁸. Figure 1b demonstrates that a sharp peak of activity is observed. The sedimentation rate of the most active fraction was previously estimated as "approximately 11S"⁷. Centrifugation for longer times in a rotor which generates greater force (SW 60Ti) collection and assay of a larger number of fractions and comparison with a wider range of markers (catalase, glyceraldehyde-3-phosphate dehydrogenase and alkaline phosphatase) reveals that this is an overestimate and that the value is nearer to 8S. Figure 1a shows that the active fractions contain a wide range of polypeptides, but that many others are removed by this simple fractionation.

When crude undiluted homogenates are heated to 80 °C for 10 min a gelatinous precipitate forms and assembly activity is lost, hence we described the activity as thermolabile³. However, when the active fractions from a sucrose gradient (or diluted, twice reclarified homogenates) are heated 10 min at 80 °C the supercoiling activity is not precipitated; it is simply made dependent on added histones and nicking-closing enzyme. Figure 1c shows that heating 10 min at 80 °C followed by cooling and centrifuging 2 min at 10,000g to remove precipitated proteins causes a dramatic simplification of the polypeptide profile. One major polypeptide (arrow in Fig. 1c) remains. It has an apparent molecular weight of 29,000, judged by comparison with known standards on SDS-gel electrophoresis (superoxide dismutase, 20,000; aldolase, 30,000; glyceraldehyde-3-phosphate dehydrogenase, 36,000; creatine kinase, 40,000; ovalbumin, 43,000). It migrates slightly slower than histone H1 (which migrates as though it has an apparent molecular weight of 27,500 in the electrophoresis conditions used here). The polypeptide of MW 29,000 is not visible in the unheated sample (Fig. 1a). It is obscured by a more prominent band which migrates just ahead of it (which is not histone H1). Note that the distribution of the polypeptide marked by an arrow in Fig. 1c is similar to the distribution of assembly activity in Fig. 1b. Most of the minor bands in the fraction marked by an arrow can be removed from the heated gradient fraction by a second centrifugation at 10,000g for 2 min. To obtain optimal precipitation of other proteins it is necessary to add 0.025 volumes of 6 mg ml⁻¹ PMSF in ethanol before heating. Ethanol alone substitutes partially but not completely.

Figure 2 shows that assembly activity also co-purifies with a polypeptide which has an apparent molecular weight of 29,000

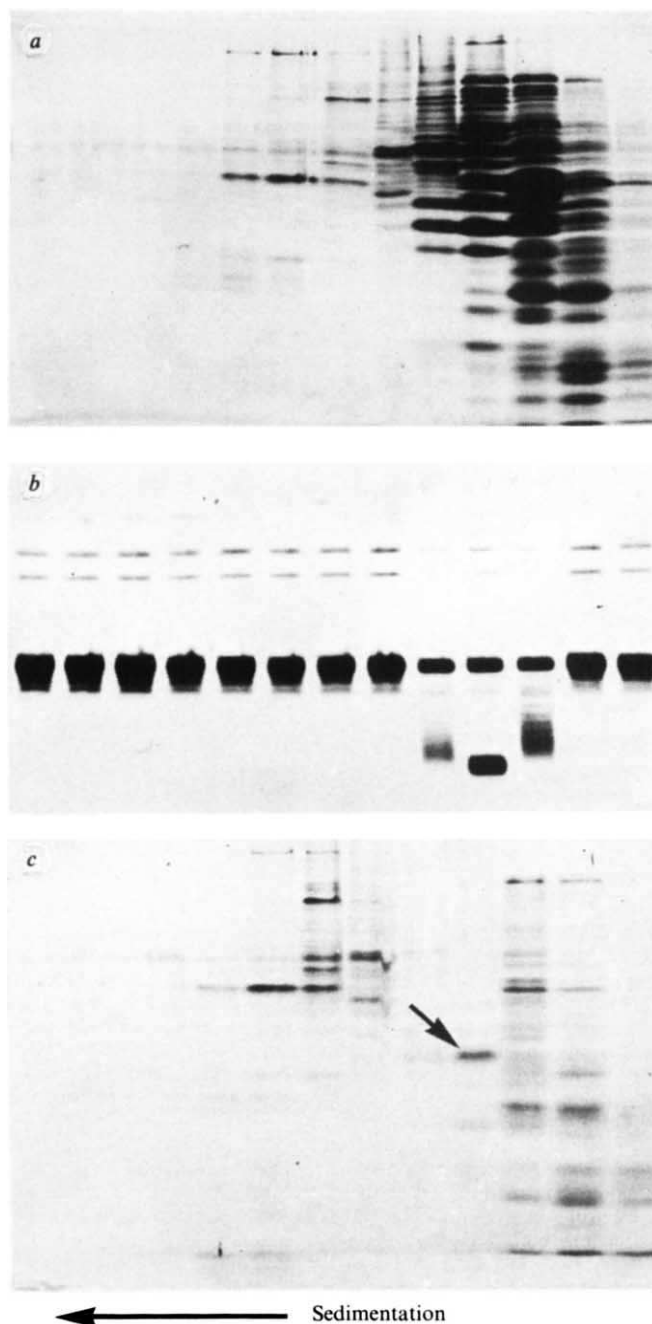


Fig. 1 Nucleosome assembly activity co-sediments with a particular protein in a sucrose gradient. *a*, SDS gel of proteins; *b*, DNA supercoiling assay; *c*, SDS gel of proteins as (*a*) which resist heat precipitation. Details are as follows: *a*, Separation of polypeptides contained in each sucrose gradient fraction by electrophoresis in a 15% polyacrylamide gel as described previously¹⁴, and visualised by Coomassie blue staining. *b*, Assay of each gradient fraction for insertion of superhelical turns into relaxed circular ³H-SV40 DNA. DNA configurations are resolved by electrophoresis through a 1% agarose gel and visualised by fluorography using 3% PPO in ethanol³. 0.1 µg of ³H-SV40 relaxed DNA was incubated in 50 µl of each gradient fraction for 3 h at 20 °C. *c*, Same as (*a*) but gradient fractions were heated 10 min at 80 °C (after addition of 0.025 volumes 6 mg ml⁻¹ PMSF in ethanol), cooled at 0 °C for 5 min, centrifuged 2 min 10,000g, at 4 °C to remove precipitated proteins and only the supernatant was treated with SDS for electrophoresis. Before gradient centrifugation 500 µl clarified egg homogenate was extracted twice with an equal volume of 1,1,2-trichlorotrifluoroethane to remove lipids and lipoproteins (such as yolk), then diluted with 100 µl of 20 mM Tris pH 7.5, 5 mM EDTA. It was centrifuged through a 10–40% sucrose gradient⁷ for 16.5 h at 38,000 r.p.m. in a Beckman SW 40Ti rotor at 4 °C. Aliquots of 50 µl were prepared for electrophoresis (*a* and *c*) and assayed for assembly (*b*).

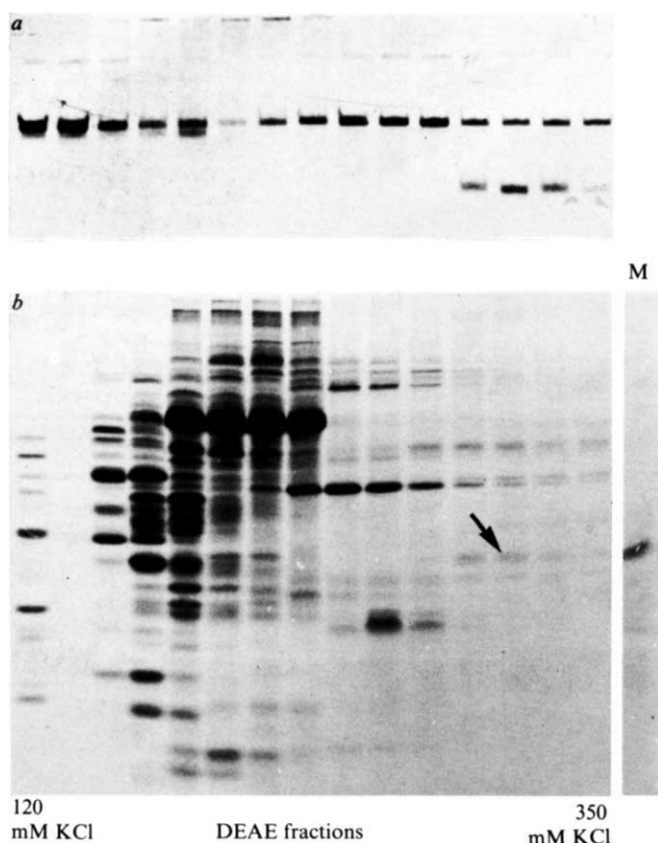


Fig. 2 Nucleosome assembly activity co-elutes with a polypeptide of MW 29,000 during anion exchange chromatography on DEAE-cellulose. *a*, Assay of each DEAE-cellulose column fraction (15 μ l diluted to 50 μ l at final KCl concentration of 120 mM) for insertion of superhelical turns into relaxed SV40 circular DNA (0.25 μ g) in the presence of calf thymus histones (0.5 μ g) and nicking-closing enzyme (see Table 1 legend). Products of a 3 h incubation at 20 °C were analysed by agarose gel electrophoresis and ethidium bromide staining. A negative image is shown to facilitate half-tone reproduction. *b*, Electrophoretic analysis of polypeptides contained in each column fraction on a 15% polyacrylamide SDS gel in conditions of Fig. 1*a* and *c*. The marker track M is the material from the track marked with an arrow in Fig. 1*c*, re-electrophoresed on this gel (that is, a heated sucrose gradient fraction). 6 ml of clarified egg homogenate was chromatographed on a 9 $\times$ 3 cm DEAE-cellulose (Whatman DE 52) column and eluted with a 150 ml linear KCl gradient from 0.12 M to 0.5 M KCl in 20 mM Tris-HCl pH 7.5. 12- μ l aliquots were diluted to 52 μ l at a final concentration of 120 mM NaCl for assay. Fractions eluting after 350 mM KCl are not shown. They contained very few proteins and showed no activity.

when clarified homogenates are fractionated by charge. Thus both the polypeptide and the activity bind tightly to the anion exchanger DEAE-cellulose suggesting that the assembly factor is highly negatively charged. After chromatography on DEAE-cellulose the assembly factor requires the addition of histones and nicking-closing enzyme in order to insert superhelical turns into relaxed DNA. When the protein of MW 29,000 is purified by heating sucrose gradient fractions and then chromatographed on DEAE-cellulose it again co-elutes with the assembly activity.

Therefore the activity which inserts superhelical turns into relaxed DNA co-purifies by both size and charge with a protein whose subunits have an apparent MW of 29,000 on SDS gels.

Figure 3*a* demonstrates the electrophoretic purity of the polypeptide prepared by sucrose gradient centrifugation followed first by heating for 10 min at 80 °C as in Fig. 1*c* and then by chromatography on DEAE-cellulose (as in Fig. 2).

Figure 3*b–d* shows that the preparation containing the purified protein illustrated in Fig. 3*a* causes the insertion of superhelical turns into relaxed DNA in the presence of acid extracted histones and nicking-closing enzyme. In the absence of the material shown in Fig. 3*a* or of the histones, assembly does not occur (Fig. 3*b–d*). Most nucleic acid contaminants are removed from the polypeptide of MW 29,000 by DEAE chromatography, but to ensure that nucleic acids or oligonucleotides are not required for its activity, a separate preparation of the protein (prepared without the final DEAE step) was incubated with 100 μ g ml⁻¹ ribonuclease A and 150 units ml⁻¹ micrococcal nuclease (in 2 mM CaCl₂) for 2 h at 25 °C followed by chromatography on Sephadex G-75 to remove oligonucleotides, CaCl₂ and the nucleases. Figure 3*e–g* illustrates that assembly activity is not decreased by this procedure.

All preparations which have demonstrated assembly activity have also contained the polypeptide of 29,000 MW. The sedimentation and gel filtration properties of the native form of this polypeptide suggest that the native protein is unlikely to consist of less than four subunits of MW 29,000.

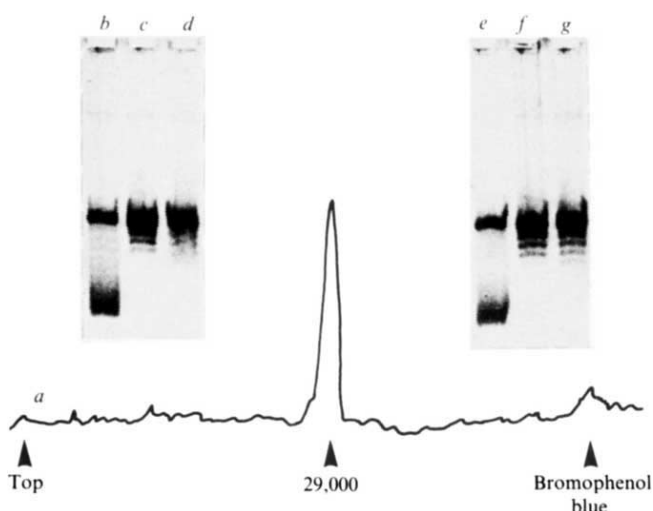
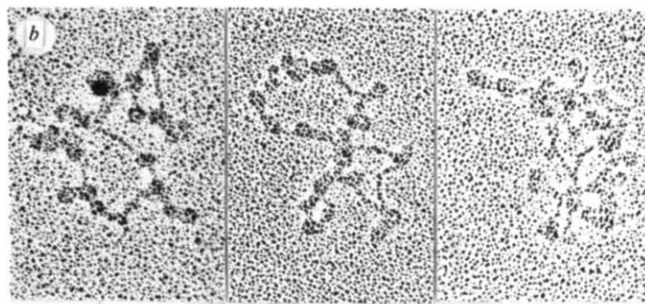
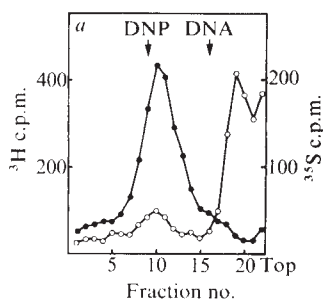


Fig. 3 The purified protein of 29,000-dalton subunits inserts superhelical turns into relaxed DNA in the presence of histones and nicking-closing enzyme. *a*, Electrophoretic purity of the polypeptide of MW 29,000 prepared by sucrose gradient centrifugation (as in Fig. 1) followed first by heating 10 min 80 °C (as in Fig. 1*c*) and then by chromatography on DEAE-cellulose (as in Fig. 2). The figure shows a microdensitometer trace of a stained 15% SDS-polyacrylamide gel. *b–d*, Assay of the material from (*a*) for insertion of superhelical turns into relaxed DNA (as in Fig. 2*a*); *b*, in the presence of histones (extracted from mouse fibroblasts by 0.2 M H₂SO₄) and nicking-closing enzyme (purified as in legend to Table 1); *c*, as *b*, but histones omitted; *d*, as *b* but the purified nucleosome assembly protein (illustrated in Fig. 3*a*) was omitted. *e–g*, as *b–d* but using assembly protein (not shown) which had been prepared by incubating a heated sucrose gradient fraction with 100 μ g ml⁻¹ ribonuclease A and 150 unit ml⁻¹ micrococcal nuclease (in 2 mM CaCl₂, 50 mM NaCl, 10 mM Tris pH 8.0) for 2 h at 25 °C followed by chromatography on Sephadex G-75 (in 0.5 mM EDTA, 50 mM NaCl, 10 mM Tris pH 7.5) to remove oligonucleotides and the nucleases.

Figure 4 illustrates two other criteria which indicate that the protein described assembles nucleosomes. The sedimentation profile of the incubation product resembles that of naturally-occurring SV40 nucleohistone complex and contains approximately equal amounts of histones and DNA (Fig. 4*a*). In addition the incubation product has the characteristic beaded appearance of the SV40 minichromosome (Fig. 4*b*).

Nicking-closing enzyme is not required for nucleosome assembly. Although it is required to relax torsion so that the insertion of superhelical turns into DNA can be visualised in the gel assay, it is not required when nucleosome assembly is

Fig. 4 The protein of 29,000-dalton subunits causes DNA and histones to form a discrete nucleoprotein complex. *a*, Velocity sedimentation of the nucleoprotein product obtained by incubating 0.1 μg ^3H -SV40 form I (55,000 c.p.m. μg^{-1}) DNA (●) with 0.2 μg ^{35}S -labelled mouse fibroblast histones (7,000 c.p.m. μg^{-1}) (○), in the presence of 20 μl of the heated peak fraction from a sucrose gradient (similar to the track marked by an arrow in Fig. 1c). After incubation for 3 h at 20 °C without nicking-closing enzyme the product was centrifuged through a 10–30% sucrose gradient in 60 mM KCl, 1 mM EDTA, 10 mM Tris pH 7.5 for 70 min at 50,000 r.p.m. in a Beckman SW60Ti rotor. The arrow marked 'DNA' indicates the position of free SV40 DNA form I and the arrow marked 'DNP' indicates the position of authentic SV40 minichromosome from infected cells. *b*, Electron micrographs of the nucleoprotein product obtained by incubation in the conditions described in *a* followed by spreading and shadowing as described previously⁷. (Nicking-closing enzyme was included in the incubations illustrated in *b*, but it is not necessary for the formation of such structures.) $\times 130,000$.



monitored by sedimentation or electron microscopy. For these assay methods the purified assembly protein (as illustrated in Fig. 3a) is sufficient to convert histones and DNA into nucleosomes.

The protein which co-purifies with assembly activity binds histones to form an active complex

Figure 5 illustrates that the purified nucleosome assembly protein binds labelled histones. All four of the nucleosome core histones bind and co-sediment as a complex with the assembly protein (at approximately 8S). The complex sediments much faster than free histones and slightly, but reproducibly, faster than the free assembly protein. Coomassie blue staining of gels followed by microdensitometry and comparison with standard histone preparations from nuclei reveals that all four of the nucleosome core histones are present in the complex in approximately equimolar amounts. Histone H-1 has not been detected so far but this could be due to technical causes.

The formation of the complex between histones and the nucleosome assembly protein is unlikely to be a nonspecific electrostatic adsorption artefact for several reasons. First, at low histone concentrations essentially all of the histones sediment as a uniform, discrete complex. Second, when the histone concentration is increased binding can be saturated without causing the complex to aggregate or sediment faster; instead the excess histones sediment as free histones. Third, the assembly protein masks sufficient positive charges on the histones to inhibit their binding to CM-cellulose at neutral pH. Fourth, if small amounts of labelled histones are added to crude egg homogenates the added histones sediment discretely as the complex described here, indicating that the vast excess of cellular proteins does not compete effectively for histone binding against the nucleosome assembly protein, even though several of the other cellular proteins are more acidic than the assembly protein. Fifth, when the complex formed between the purified assembly protein and labelled histones is isolated from a sucrose gradient and incubated with DNA the histones are transferred to the DNA with the formation of a beaded nucleoprotein complex which contains supercoiled DNA.

Cross-linking studies should provide further information on the specificity and stoichiometry of histone binding.

Properties of the nucleosome assembly protein

The protein which co-purifies with nucleosome assembly activity electrophoreses as a rather broad band with an apparent MW of 29,000 in SDS-polyacrylamide gels. When it is electrophoresed over long distances it resolves into two distinct bands, suggesting that the band seen in Figs 1–3 probably

contains more than one polypeptide chain. Its sedimentation and gel filtration properties suggest that the native protein is larger than 100,000 daltons. It does not bind to CM-cellulose above pH 6 but binds to DEAE-cellulose and elutes between 0.3 M and 0.4 M KCl at pH 7.5, suggesting that it is strongly

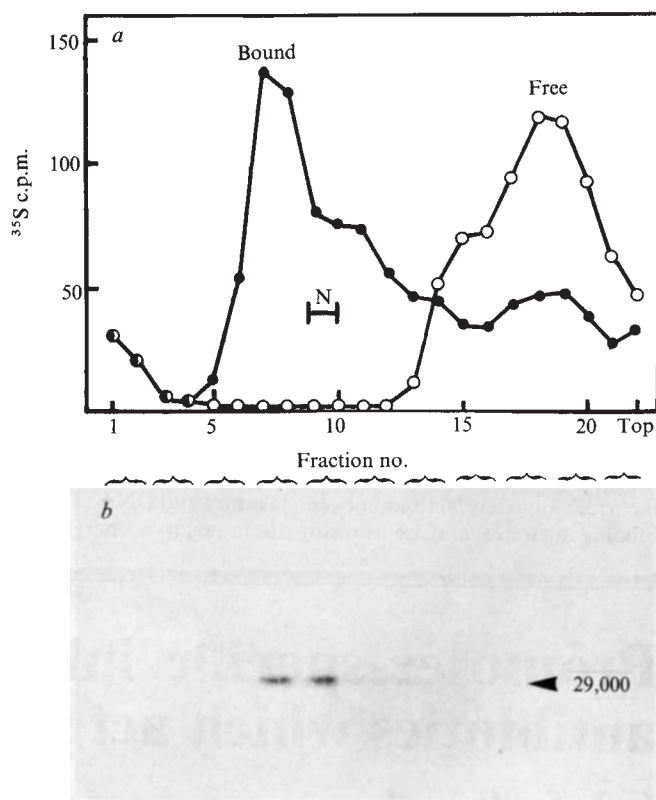


Fig. 5 The nucleosome assembly protein binds histones. *a*, distribution of ^{35}S -methionine labelled, acid-extracted mouse histones after centrifugation through a 10–30% sucrose gradient. ●, Histones were incubated with a 10 $\times$ concentrated (Minicon B-15) heated sucrose gradient fraction (as marked by arrow in Fig. 1c) with 50 mM NaCl for 20 min before centrifugation in a separate tube from the free histone marker (○). Radioactivity was measured by liquid scintillation counting 50- μl aliquots of each fraction in Aquasol. Fractions from (*a*) (●) were pooled in pairs and 50- μl aliquots were electrophoresed in a 15% polyacrylamide gel (as in Fig. 1a) to locate the polypeptide of MW 29,000 by Coomassie blue staining. (At the histone concentration used in this figure histone bands were stained too faintly for photographic reproduction.) The bracket marked N indicates the position of the nucleosome assembly protein when it is centrifuged in the absence of histones (fractions 9 and 10). Centrifugation was for 20.5 h at 50,000 r.p.m. in a Beckman SW60Ti rotor at 4 °C. Gradients contained in 2mM EDTA, 20mM Tris-HCl pH 7.5.

acidic. This conclusion is confirmed by isoelectric focusing. Thus on two dimensional gels¹¹ the protein has an isoelectric point of 5.

Additional evidence that this protein is indeed responsible for nucleosome assembly comes from the observation that it shows the same remarkable resistance to thermal denaturation and proteolytic degradation as the assembly activity. Thus its sedimentation as a multi-subunit structure and its ability to bind histones are not affected by heating. After prolonged treatment with chymotrypsin the protein's sedimentation constant is only slightly decreased and it continues to bind histones. Electrophoresis reveals that at least one cut has occurred in each polypeptide chain, since the band of MW 29,000 is replaced by two large fragments of MW approximately 26,000 but which still sediment as a multi-subunit structure. Proteinase K at pH 10 which abolishes the assembly activity also cleaves the protein, generating fragments of less than 20,000 daltons.

Five ml of egg homogenate (3 ml packed eggs + 2 ml buffer) containing 1.25 g total protein yields 160 µg of pure protein. 1 µg of pure protein can convert at least 0.25 µg DNA into nucleosomes in the presence of 0.5 µg histones.

The nucleosome assembly process

When histones and DNA are mixed in solution at physiological ionic strength, they instantly form a precipitate. When the protein described here is mixed with histones before addition of DNA, precipitation is prevented and nucleosomes form. We observed previously that the egg's histone pool is complexed to a large negatively charged molecule and that fractions containing this complex interact with DNA to form nucleosomes⁷. Therefore we proposed that histones are organised into nucleosome precursor complexes by an acidic factor which has a 'catalytic' role in the assembly process⁷. The protein we have purified clearly binds histones and transfers them to DNA to form nucleosomes. The possibility that it is recycled remains to be tested.

Nucleosomes can be assembled by prolonged dialysis from 2 M NaCl (reviewed in ref. 2). Therefore it seems that the spatial information required for histone-histone interactions and histone-DNA interactions resides in the histones and DNA themselves. We suggest that the role of the protein we have purified is that of a 'molecular chaperone' which prevents incorrect ionic interactions between histones and DNA. Thus by binding histones and neutralising their positive charges the

nucleosome assembly protein could prevent nonspecific ionic interactions and allow only selected specific interactions to occur. The additional possibility that specific interactions between different histones occur in the complex formed with the nucleosome assembly protein remains to be tested.

The nucleosome assembly reaction studied here seems to have very simple requirements. It requires only the protein preparation illustrated in Fig. 3a, DNA and acid extracted (or salt extracted) histones incubated in NaCl (or KCl) concentrations between 50 mM and 150 mM. It is inhibited by higher NaCl concentrations. The reaction will proceed to completion in 3 h at 0 °C. We have been unable to find requirements for low molecular weight co-factors, such as Mg²⁺ or ATP.

Eggs of *Xenopus laevis* are unusual in that they divide to form 5,000 cells in 5 h with some cell cycles of only 10 min. In addition they accumulate a large histone pool, sufficient for 15,000 nuclei⁴⁻⁶. Therefore it is possible that the nucleosome assembly protein we have described is only an adaptation to this rapid rate of chromosomal replication. However, during the course of this work we have found that mouse fibroblast nuclei and wheat embryos also contain a thermostable protein which converts histones and DNA into nucleosomes (B.M.H., A.D.M. & R.A.L., in preparation). Therefore it will be interesting to determine if the protein we have described is a universal mechanism for assembling the nucleosome subunits of eukaryotic chromatin.

We thank John Gurdon, Aaron Klug and Doug Melton for comments on the manuscript, and Geoff Partington for suggesting extraction with trichloroethane. B.M.H. is a post-doctoral fellow of the Canadian MRC.

Received 13 June; accepted 17 August 1978.

1. Kornberg, R. D. *A. Rev. Biochem.* **46**, 931-954 (1977).
2. Felsenfeld, G. *Nature* **271**, 115-122 (1978).
3. Laskey, R. A., Mills, A. D. & Morris, N. R. *Cell* **10**, 237-243 (1977).
4. Adamson, E. D. & Woodland, H. R., *J. molec. Biol.* **88**, 263-285 (1974).
5. Woodland, H. R. & Adamson, E. D. *Dev. Biol.* **57**, 118-135 (1977).
6. Adamson, E. D. & Woodland, H. R. *Dev. Biol.* **57**, 136-149 (1977).
7. Laskey, R. A. *et al. Cold Spring Harb. Symp. quant. Biol.* **42**, 171-178 (1977).
8. Germond, J. E., Hirt, B., Oudet, P., Gross-Bellard, M. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1843-1847 (1975).
9. Gellert, M., Mizuuchi, K., O'Dea, M. H. & Nash, H. A. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3872-3876 (1976).
10. Ebeling, W. *et al. Eur. J. Biochem.* **47**, 91-97 (1974).
11. O'Farrell, P. H. *J. biol. Chem.* **250**, 4007-4021 (1975).
12. Mattoccia, E., Gandini-Attardi, D. & Tocchini-Valentini, G. P. *Proc. natn. Acad. Sci. U.S.A.* **73**, 4551-4554 (1976).
13. Tegtmeyer, P. & Macasaet, F. *J. Virol.* **10**, 599-604 (1972).
14. Laskey, R. A., Gurdon, J. B. & Crawford, L. V. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3665-3669 (1972).

Promoter-specific inhibition of transcription by antibiotics which act on DNA gyrase

C. L. Smith*

The Public Health Research Institute of The City of New York, Inc., 455 First Avenue, New York, New York 10016

M. Kubo & F. Imamoto

Research Institute for Microbial Diseases, Osaka University, Yamada-Kami, Suita, Osaka, Japan

Nalidixic acid, novobiocin and coumermycin specifically inhibit phage promoter-dependent transcription of the trp operon in $\phi 80$ ptp but not transcription from the authentic trp promoter. The nalidixic acid inhibition is not observed in an E. coli strain containing a nalA^r mutation. These results indicate that DNA gyrase is involved in transcription.

* Present address: Department of Biological Sciences, Columbia University, New York, New York 10027.

DNA GYRASE catalyses the ATP dependent introduction of negative superhelical twists into double stranded closed circular DNA in *Escherichia coli*¹. This activity is required for λ integrative recombination² and replication of *E. coli*^{3,4}, plasmids^{5,6} and bacteriophages⁷⁻⁹. DNA gyrase is inhibited by the antibiotics coumermycin and novobiocin¹⁰. The *cou* locus mapping close to the *dnaA* at 82 min on the *E. coli* map determines resistance to these drugs^{11,12}.

DNA gyrase activity is also inhibited by nalidixic acid (NAL) and oxolinic acid^{13,14}. Resistance to these drugs is determined by

the *nalA* locus mapping at 48 min on the *E. coli* map¹⁵. It is thought that the nick closing activity required in the supercoiling reaction resides in the NAL target protein, *pnaI*, because (1) the ATP independent DNA gyrase catalysed relaxation of supercoiled DNA is inhibited by NAL and oxolinic acid but not by novobiocin or coumermycin and (2) incubation of DNA gyrase and DNA with NAL or oxolinic acid results in the formation of a relaxation type complex.

At low doses NAL is thought to be a specific inhibitor of DNA replication in *E. coli*. However at higher doses several workers have observed inhibition of transcription¹⁶⁻¹⁸. In preliminary experiments we observed indications that NAL interfered with $\phi 80$ phages transcription at doses normally thought to inhibit only DNA replication. Here we report the effect of NAL on transcription of the tryptophan (*trp*) operon integrated into the $\phi 80$ phage genome.

Nalidixic acid is a promoter-specific inhibitor of *trp* expression

When the bacterial *trp* operon is integrated into the λ phage genome, transcription of the operon may initiate at either the authentic *trp* promoter (P_{trp}) or at the leftward phage promoter as shown in Fig. 1^{19,20}. Initiation of transcription from P_{trp} is turned off in the presence of tryptophan and the *trp* repressor. Read-through transcription of the *trp* operon from the leftward phage promoter can occur in the presence of the *trp* repressor but is completely inhibited when phage repressor molecules are present. Subsequent studies yielded identical results with $\phi 80$ phages^{21,22}.

Infection of a nonlysogenic *E. coli trpE*⁻ strain (MO1512) with bacteriophage $\phi 80$ ptrp, in the presence of tryptophan, led to immediate synthesis of anthranilate synthetase (ASase; product of the *trpED* genes) as shown in Fig. 2a. ASase synthesis occurred by the read-through transcription of the *trp* operon since no phage repressor molecules were present in the host cell. However, when the $\phi 80$ ptrp infected cells were incubated in the presence of 40 $\mu\text{g ml}^{-1}$ NAL, no ASase activity was detected even after prolonged incubation.

The effect of NAL on ASase synthesis originating at the P_{trp} promoter is shown in Fig. 2b. In this experiment *E. coli* strain MO1513 *trpE*⁻ ($\phi 80^+$) was infected with $\phi 80$ ptrp in the absence of tryptophan. ASase synthesis, due to transcription from the authentic *trp* promoter, began immediately after removal of tryptophan from the culture medium, as shown by the control cells in Fig. 2b. When NAL was present only a slight decrease in ASase activity was seen throughout the incubation period. Dose-response curves for a 60-min incubation are shown in Fig. 3a. Concentrations of NAL which virtually eliminate ASase synthesis originating from the phage promoter have essentially no effect on synthesis originating from P_{trp} .

Nalidixic acid inhibition occurs in the absence of DNA replication

The amount of ASase synthesised in the procedure used in Figs 2 and 3 is regulated by the number of $\phi 80$ ptrp templates available for transcription²². Thus, it was possible that NAL inhibited ASase synthesis simply by interfering with phage replication. Therefore, the effect of NAL on phage promoter dependent ASase synthesis was examined in the absence of DNA replication. An *E. coli* strain lacking DNA polymerase I and II and containing a temperature-sensitive DNA polymerase III (*polA*⁻, *polB*⁻, *polC*_{ts}) was infected with $\phi 80$ ptrp and treated with NAL. As shown in Fig. 4a, addition of NAL to $\phi 80$ ptrp infected cells incubated at 30°C led to rapid partial inhibition of ASase synthesis by the read-through mechanism. Addition of NAL to $\phi 80$ ptrp-infected cells incubated at the restrictive temperature also led to inhibition of phage promoter controlled ASase synthesis (Fig. 4b). However, somewhat less inhibition was observed at the restrictive temperature than at the permis-

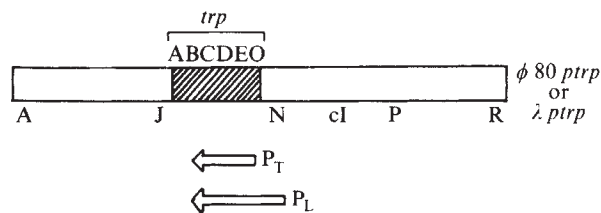


Fig. 1 Transcriptional map of bacteriophages $\phi 80$ and λ carrying the *E. coli trp* operon. P_T , authentic *trp* promoter (repressed by tryptophan but not by phage repressor); P_L , early leftward phage promoter (repressed by the phage repressor but not by tryptophan).

sive temperature. This indicates a small portion of the NAL inhibition may arise from interference with replication but most of the effect is replication independent.

Nalidixic acid inhibits phage promoter-directed *trp* transcription

DNA-RNA hybridisation experiments were conducted to determine whether NAL inhibition was occurring at the level of transcription or translation. The results of these experiments are summarised in Table 1. Addition of NAL to $\phi 80$ ptrp infected cells reduced the amount of phage promoter-controlled *trp* mRNA synthesis to about 10% of the level observed in the absence of the drug. On the other hand, the drug caused a slight increase in the amount of mRNA derived from the P_{trp} promoter. The effects observed with chloramphenicol are very different. Chloramphenicol specifically inhibits P_{trp} transcription but does not affect phage transcription. Similar chloramphenicol effects have been observed in λ ptrp phages^{23,24}.

We considered the possibility that NAL was stimulating degradation of the phage promoter initiated mRNA rather than inhibiting mRNA synthesis. The effect of NAL on the stability of preformed mRNA is shown in Fig. 5. In these experiments rifampicin was added at the time of addition of NAL to inhibit mRNA initiation. NAL did not significantly increase the rate of *trp* mRNA degradation.

At low concentrations NAL inhibits DNA replication^{18,25,26}. At these concentrations transcription is not inhibited as judged both by incorporation of labelled precursor into RNA^{18,25,26} and by direct examination of several bacterial operons¹⁷. Based on

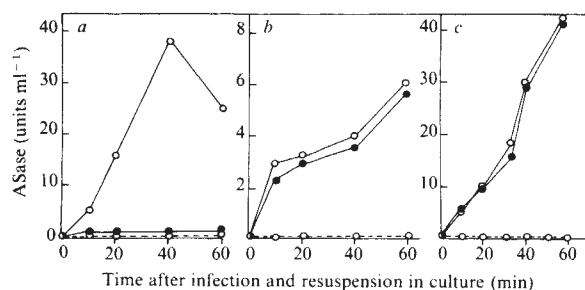


Fig. 2 Nalidixic acid inhibition of ASase synthesis directed by a $\phi 80$ ptrp template. (a) Read-through transcription from the phage promoter in a *nalA*⁺ strain. In the presence of tryptophan, strain MO1512 *trpE*⁻ was grown to a density of 2×10^8 cells ml^{-1} , infected with $\phi 80$ ptrp and treated with 40 $\mu\text{g ml}^{-1}$ NAL (●). (b) Transcription from the authentic *trp* promoter in a *nalA*⁺ strain. In the absence of tryptophan, strain MO1513 *trpE*⁻ ($\phi 80^+$) was grown to a density of 2×10^8 cells ml^{-1} and in the absence of tryptophan, infected with $\phi 80$ ptrp and treated with 40 $\mu\text{g ml}^{-1}$ NAL (●). (c) Read-through transcription from the phage promoter in a *nalA*⁺ strain. In the presence of tryptophan strain MO1512 *trpE*⁻ *nalA*⁺ was grown to a density of 2×10^8 cells ml^{-1} , infected with $\phi 80$ ptrp and treated with 40 $\mu\text{g ml}^{-1}$ NAL (●). Controls: in all cases were—infected with $\phi 80$ ptrp but not treated with NAL (○—○); not infected with $\phi 80$ ptrp and not treated with NAL (○—○). Experimental procedures are described in full elsewhere²³.

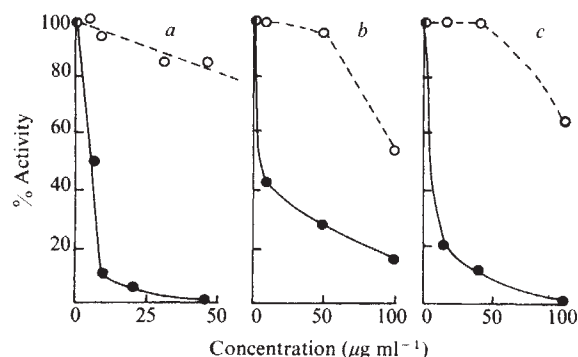


Fig. 3 Inhibition of ASase synthesis by various concentrations of nalidixic acid (a), novobiocin (b) and coumermycin (c). The per cent activity was calculated from the amount of ASase activity observed at 60 min at the indicated dose. 100% activity was 39 units ml⁻¹ for phage directed synthesis (●) and 6.7 units ml⁻¹ for P_{trp} directed synthesis (○). Experimental procedures are described in Fig. 2.

these results it has been assumed that NAL is a specific inhibitor of DNA replication. However, our results show that transcription monitored by the production of competent mRNA and also by enzyme synthesis can be inhibited by the same NAL doses previously believed only to inhibit DNA replication. Therefore, NAL should not be used as a specific inhibitor of DNA replication.

In 1973, Puga and Tessman¹⁶ reported that nalidixic acid interfered with plasmid phage S13 mRNA synthesis at drug levels which did not affect host mRNA synthesis. Since all S13 mRNA synthesis was equally inhibited and it had been previously observed that intercalating dyes inhibited transcription²⁷, these authors proposed that the specificity of the inhibition was due to a preferential intercalation of the antibiotic into the supercoiled plasmid DNA. However, our results show that NAL inhibits, with total specificity, one of the two promoters both of which serve the same operon and are on the same

plasmid molecule. Therefore, the simple notion that superhelical density generates NAL specificity must be rejected.

In experiments shown in Fig. 4 only partial inhibition of phage promoter controlled ASase synthesis was observed. This is in contrast to the complete inhibition observed in the experiment shown in Fig. 2. In the former experiment NAL was added after initiation of transcription whereas in the latter experiment it was added just before. However, the former experiment also involved a multiple mutant strain while the latter involved a wild-type strain. To determine which of these factors was significant NAL was added to the same wild-type strain used for Fig. 2 but after transcription was initiated. A partial inhibition of ASase synthesis was seen just like that shown in Fig. 4 (data not shown). As the effect of NAL is greater when added before transcription initiates, it is likely that NAL acts preferentially on initiation.

DNA gyrase is involved in transcription

Mutations at the *nalA* locus confer NAL resistance to DNA synthesis^{18,27}. To see if this same locus was involved in NAL sensitive transcription we examined ASase synthesis in a derivative of *E. coli* strain MO1512 containing the *nalA26* mutation (Fig. 2c). No NAL inhibition of phage promoter directed synthesis was observed in the *nalA*^r *E. coli* strain in conditions where total inhibition is seen in the isogenic *nalA*^s strain (Fig. 2a). This shows the NAL-target protein involved in transcription is the same protein involved in DNA replication.

The recently isolated *nalA* protein participates in DNA gyrase mediated reactions^{1,10,13,14}. Since this enzyme is also inhibited by coumermycin and novobiocin, we examined the

Table 1 Nalidixic acid inhibition of *trp* mRNA synthesis

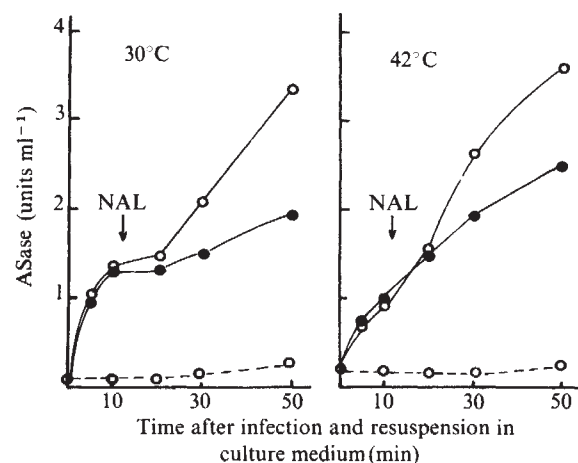
	Anti-biotic	Tryptophan (50 μg ml ⁻¹)	<i>trp</i> mRNA (% of total RNA)	Specific activity of ³ H-RNA (c.p.m. per mg)
Phage promoter directed synthesis (strain AE1)	None	+	1.17	5.8 × 10 ³
	NAL	+	0.12	6.3 × 10 ³
	CAP	+	3.01	5.6 × 10 ³
P _{trp} directed synthesis (strain AE1 (h ^Δ 1 ⁸⁰))	None	—	3.02	3.3 × 10 ³
	NAL	—	3.48	1.3 × 10 ³
	CAP	—	0.18	3.0 × 10 ³

E. coli strain *trpAE1* or its lysogen *trpAE1* (h^Δ1⁸⁰) was infected with ϕ 80*trp* phages and incubated at 30 °C in the presence or absence of L-tryptophan (50 μg ml⁻¹), respectively. NAL (40 μg ml⁻¹) or chloramphenicol (CAP; 100 μg ml⁻¹) was added at the 8th minute of incubation after the phage infection. Pulse-labelling was carried out with 100 μCi of ³H-uridine for 1 min starting 29.5 min after infection. The value of *trp* mRNA hybridised with λ *trp* DNA is expressed as a percentage of total labelled RNA added into a hybrid assay system and is the average of duplicate determinations. The value of λ DNA background was subtracted from each hybrid value. Details for procedures used in these experiments are described elsewhere³⁹.

effect of these antibiotics on ASase synthesis. The inhibition of P_{trp} and phage promoter directed ASase synthesis by different doses of novobiocin and coumermycin (Fig. 3b, c) show a striking parallel with NAL inhibition. In all cases P_{trp} directed synthesis is much less sensitive to these antibiotics than phage promoter directed ASase synthesis. The 20% inhibition of P_{trp} directed ASase synthesis at 100 μg ml⁻¹ novobiocin probably reflects inhibition of DNA replication observed at this dose level in this strain (see ref. 28). The greater inhibition observed with coumermycin could be due to the greater permeability of *E. coli* to this antibiotic¹¹.

Our results provide the first direct evidence that DNA gyrase is involved in transcription. Three other antibiotics known to be specific inhibitors of DNA gyrase mediated supercoiling dis-

Fig. 4 Nalidixic acid inhibition of phage promoter dependent ASase synthesis in the presence (30 °C) and absence (42 °C) of DNA replication. Strain BT10266 *polA*⁻, *polB*⁻, *polC*₁₈ was grown to 2 × 10⁸ cells ml⁻¹ by incubation at 30 °C in minimal medium containing tryptophan²³. Following infection with ϕ 80*trp* the culture was divided into two portions, one of which was further incubated in fresh minimal medium at 30 °C while the other was incubated at 42 °C. After 15 min, 40 μg ml⁻¹ NAL was added to portions of each of the cultures and the incubation at the respective temperature continued (●—●). Controls: infected with ϕ 80*trp* but not treated with NAL (○—○); not infected with ϕ 80*trp* and not treated with NAL (○—○).



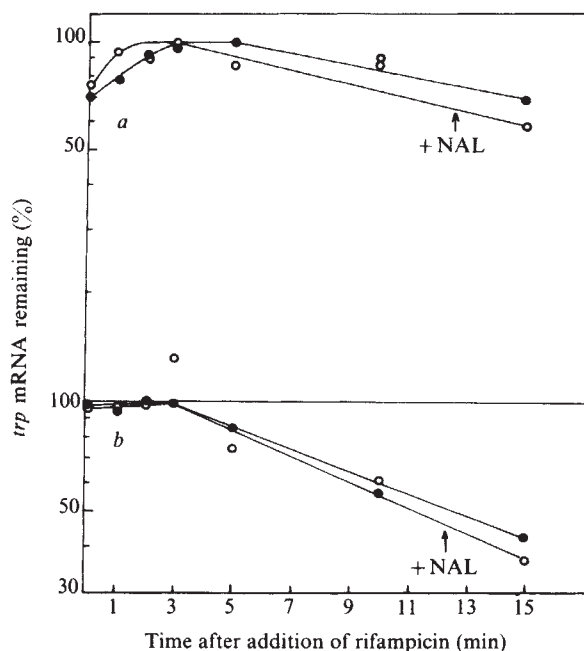


Fig. 5 Effect of nalidixic acid on the stability of *trp* mRNA synthesised from a phage promoter (a) and the authentic *trp* promoter (b). At 30 °C strain *trpAEl* (a) or strain *trpAEl* ($h^+ i^{80}$) (b) was infected with $\phi 80trp$ in the presence of (a) or absence of (b) L-tryptophan ($50 \mu\text{g ml}^{-1}$), respectively. Pulse-labelling (1 min) carried out with $33 \mu\text{Ci ml}^{-1}$ ^3H -uridine at the 15th minute after the phage infection was immediately followed by addition of rifampicin ($300 \mu\text{g ml}^{-1}$) in the presence of (○) or absence (●) of $40 \mu\text{g ml}^{-1}$ NAL. At the times indicated portions of the cultures (6×10^9 cells) were removed for RNA extraction. The *trp* mRNA values are expressed as ^3H -RNA hybridised per μg RNA and normalised to 100% for the maximum value. The value represented is the average of duplicate determinations. Detailed procedures are described elsewhere³⁹.

played the same promoter specific inhibition of transcription as NAL. The specificity implies that transcription requires DNA gyrase when initiation occurs at the phage promoter but not when initiation occurs at the *trp* promoter. In addition, it seems that it is the supercoiling activity of DNA gyrase that is involved in transcription. The effectiveness of these inhibitors on phage promoter-directed transcription is striking when compared to the inhibitor concentration required to obtain an equivalent degree of inhibition of DNA synthesis.

The origin of the specific requirement of DNA gyrase for one of the two promoters remains unexplained. Botcham²⁹ observed an increase in early phage promoter mRNA synthesis in λ when supercoiled DNA template was used instead of a relaxed template. Although no studies have directly compared *P_{trp}* transcription on supercoiled and relaxed templates, it is possible that by raising the superhelix density DNA gyrase could preferentially assist a weak promoter at binding RNA polymerase. This simple mechanism would be consistent with suggestions that complex formation between RNA polymerase and a promoter is accompanied by duplex unwinding^{30,31}. However, the DNA gyrase-requiring promoter seems to be stronger than the *P_{trp}* promoter when total enzyme activity is examined. It is difficult to explain this simply by considerations of superhelix density.

As an alternative it is tempting to speculate that the critical factor in determining the requirement for DNA gyrase is the linking of transcription with translation^{23,24}. The phage promoter which requires DNA gyrase is not linked to translation whereas the gyrase-independent promoter is linked to translation. Clearly, further studies are needed to determine whether

it extends to other host cell translational-independent promoters.

Another alternative hypothesis is that the phage promoter is subject to some type of topological constraint which DNA gyrase must overcome in order for transcription to occur. For example the phage promoter could be involved in a loop of DNA where the ends were held rigid by attachment to a cell constituent such as the cell membrane. Such a loop structure had been postulated for the *E. coli* chromosome^{32,33}.

In λ phage it has been reported that transcription initiated at the leftward phage promoter is dependent on the phage N protein³⁴⁻³⁷. This protein interacts with RNA polymerase and DNA so that in the presence of the rho termination factor, the RNA polymerase no longer recognises a termination point present in the phage genome³⁸. Transcription of the *trp* operon in $\phi 80$, is dependent on the phage N gene product, as the operon is integrated into the phage at a point past the rho termination site. Therefore, it is possible that NAL is acting as an anti-N effector, so that phage transcription in the presence of the drug is terminated at the rho site. However, in preliminary experiments, NAL inhibition of phage controlled ASase synthesis has been observed in λtrp phages lacking the rho dependent termination site (data not shown).

Although it is not clear how *pnaI* participates in RNA synthesis, it is obvious that the transcriptional apparatus at some operons is dependent on the *pnaI* protein. Thus, NAL and *pnaI* can serve as a probe to examine the transcriptional process at least in the case of phage promoter synthesis and perhaps in bacterial transcription.

We thank Dr Michio Oishi for helpful discussions and Ms J. Tamayo for preparing the manuscript. This work was supported by a grant from NIH (GM21073) to Michio Oishi. C.L.S. is the recipient of a PHS Postdoctoral Fellowship (GM95460).

Received 16 May; accepted 31 July 1978.

- Gellert, M., Mizuuchi, K., O'Dea, M. & Nash, H. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3872-3876 (1976).
- Mizuuchi, K. & Nash, H. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3524-3528 (1976).
- Crumplin, G. C. & Smith, J. T. *Nature* **260**, 643-645 (1976).
- Pisetsky, D., Berkower, I., Wickner, R. & Hurwitz, J. *J. molec. Biol.* **71**, 557-571 (1972).
- Sakakibara, Y. & Tomizawa, J. *Proc. natn. Acad. Sci. U.S.A.* **71**, 802-806 (1974).
- Staudenbauer, W. L. *Molec. gen. Genet.* **145**, 273-280 (1976).
- Itoh, I. & Tomizawa, J. *Nature* **270**, 78-80 (1977).
- Marians, K. J., Ikeda, J., Schlagman, S. & Hurwitz, J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1965-1968 (1977).
- Schneck, P. K., Staudenbauer, W. L. & Hofschneider, P. H. *Eur. J. Biochem.* **38**, 130-136 (1973).
- Gellert, M., O'Dea, M., Itoh, T. & Tomizawa, J. *Proc. natn. Acad. Sci. U.S.A.* **73**, 4474-4478 (1976).
- Ryan, M. J. *Biochemistry* **15**, 3769-3777 (1976).
- Ryan, M. J. & Wells, R. D. *Biochemistry* **15**, 3778-3782 (1976).
- Sugino, A., Peebles, C. L., Kreuzer, K. N. & Cozzarelli, N. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4767-4771 (1977).
- Gellert, M., Mizuuchi, K., O'Dea, M., Itoh, T. & Tomizawa, J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4772-4776 (1977).
- Hane, M. & Wood, T. *J. Bact.* **99**, 238-241 (1969).
- Puga, A. & Tessman, I. *J. molec. Biol.* **75**, 99-108 (1973).
- Schuman, H. & Schwartz, M. *Biochim. biophys. Res. Commun.* **64**, 204-209 (1975).
- Staudenbauer, W. L. *Eur. J. Biochem.* **62**, 491-497 (1976).
- Ihara, S. & Imamoto, F. *Biochim. biophys. Acta* **432**, 199-211 (1976).
- Nishimune, Y. *Virology* **53**, 236-246 (1973).
- Inoko, H., Naito, S., Ito, K. & Imai, M. *Molec. gen. Genet.* **129**, 37-48 (1974).
- Smith, C. L. & Oishi, M. *Molec. gen. Genet.* **148**, 131-138 (1976).
- Imamoto, F. & Tani, S. *Nature new Biol.* **240**, 182-175 (1972).
- Imamoto, F. *J. molec. Biol.* **43**, 51-69 (1969).
- Goss, W. A., Deitz, W. H. & Cook, T. M. *J. Bact.* **89**, 1068-1074 (1965).
- Goss, W. A., Deitz, W. H. & Cook, T. M. *J. Bact.* **88**, 1112-1118 (1964).
- Hurwitz, J. & Rosano, C. L. *Biochim. biophys. Acta* **108**, 697-699 (1965).
- Smith, C. L. & Oishi, M. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1657-1661 (1978).
- Botchan, P. *J. molec. Biol.* **105**, 161-176 (1976).
- Saucier, J. M. & Wang, J. C. *Nature new Biol.* **239**, 167-170 (1972).
- Chamberlin, M. A. *Rev. Biochem.* **43**, 721-776 (1974).
- Worcel, A., Burgi, E., Robinton, J. & Carlson, C. L. *Cold Spring Harb. Symp. quant. Biol.* **38**, 43-52 (1973).
- Pettijohn, D. E. & Hecht, R. *Cold Spring Harb. Symp. quant. Biol.* **38**, 31-42 (1973).
- Franklin, N. in *The Bacteriophage Lambda* (ed. Hershey, A. D.) 621-638 (Cold Spring Harbor Laboratory, New York, 1971).
- Franklin, N. C. *J. molec. Biol.* **89**, 33-48 (1974).
- Adhya, S., Gottesman, M. & Crombrugge, B. de *Proc. natn. Acad. Sci. U.S.A.* **71**, 2534-2538 (1974).
- Nakamura, H., Yasunobu, K. & Imamoto, F. *Molec. gen. Genet.* **159**, 13-20 (1978).
- Roberts, J. *Cold Spring Harb. Symp. quant. Biol.* **35**, 121-133 (1970).
- Yamamoto, T. & Imamoto, F. *J. molec. Biol.* **92**, 289-309 (1975).

Bacteriophages inhibit degradation of abnormal proteins in *E. coli*

Lee D. Simon & Kathleen Tomczak

Waksman Institute of Microbiology

Ann C. St. John

Department of Microbiology, Rutgers, The State University, Piscataway, New Jersey 08854

On infection of Escherichia coli cells by bacteriophages T4, T5 or T7, the degradation of E. coli protein fragments and abnormal proteins is inhibited. Normal E. coli proteins, however, continue to be degraded at their usual rates. T4 early protein(s) is needed to inhibit the turnover of abnormal proteins in T4-infected E. coli cells.

HALF LIVES of proteins with abnormal conformations and protein fragments are much shorter than those of most normal proteins in bacterial and in eukaryotic cells^{1,2}. Most nonsense fragments of *Escherichia coli* proteins are so short-lived that either they fail to accumulate to detectable levels or they are found in much lower quantities than wild-type proteins^{3,4}. *E. coli* polypeptides containing missense alterations also tend to turn over rapidly compared to normal proteins⁵. Nonsense fragments of virus proteins from bacteriophage T4-infected *E. coli* cells are, in contrast, often found at levels approximating those of normal T4 proteins⁶. In this article we present evidence relating to the basis for this difference in stability between *E. coli* and T4 protein fragments.

The identities and functions of the specific enzymes responsible for the rapid turnover of abnormal proteins and protein fragments are not known in detail, but some information of interest is available about this system in *E. coli*. The degradation of such proteins is ATP-dependent². If cellular ATP levels are substantially reduced, the turnover of abnormal proteins and protein fragments is also reduced⁷⁻⁹. To investigate the characteristics of this degradative system, the intracellular levels of abnormal proteins can be increased by treating cells with puromycin, leading to the production of protein fragments, or with amino acid analogues, leading to the production of full-sized proteins with abnormal conformations. In *E. coli*, both puromycyl peptides and analogue-containing proteins are degraded by an ATP-dependent system^{2,7,8}, as are other abnormal proteins and protein fragments. Specific mutations called *deg*⁻ (or *lon*⁻) at 11.4 minutes on the *E. coli* map, near *proC*, reduce the turnover rates of protein fragments¹⁰, abnormal proteins, and also of phage λ proteins containing temperature-sensitive, missense alterations¹¹. The molecular basis for the effect of *deg*⁻ mutations on the *E. coli* degradation system is obscure.

The mechanism by which the *E. coli* degradation system recognises abnormal proteins is also obscure. Electron micrographs suggest that abnormal proteins, if present in sufficient quantity, form intracellular aggregates which appear adjacent to the *E. coli* inner membrane^{12,13}. The ATP-dependent degradative activity is probably bound to the *E. coli* cell membrane (R. Voellmy and A. L. Goldberg, personal communication). The adherence of abnormal proteins to the cell membrane, probably through exposed hydrophobic regions of the proteins, may provide the recognition signal for the degradation system. It may be of interest, therefore, that the earliest detectable step in T4 capsid assembly involves the formation of amorphous aggregates of unprocessed capsid proteins on the cell membrane¹⁴. These aggregates are morphologically indistinguishable from those formed by abnormal *E. coli* proteins¹. Yet, if these T4

proteins were degraded, the production of progeny virions could not occur. The mechanism by which T4 proteins probably avoid the *E. coli* degradation system is described below.

Degradation of *E. coli* proteins and puromycyl peptides after T4 infection

To study the effects of bacteriophage infection on the degradation of abnormal *E. coli* proteins and protein fragments, we first examined the degradation of puromycyl protein fragments (PPF) in uninfected and in T4-infected *E. coli* CR63 cultures. We exposed a CR63 culture to ¹⁴C-leucine in the presence of puromycin; the unincorporated puromycin and ¹⁴C-leucine were then removed and the cells were resuspended in fresh medium with unlabelled leucine. The culture containing the ¹⁴C-PPF was divided into two equal volumes. T4 particles in tryptone broth were added to one volume of cells and sterile broth was added to the other. Samples taken before the culture was divided and at various times thereafter were added to cold trichloroacetic acid (TCA) at a final TCA concentration of 5% to precipitate high molecular weight polypeptides. As ¹⁴C-PPF are degraded to amino acids which are soluble in TCA the corresponding radioactivity appears in the TCA supernatant. The fraction of the ¹⁴C-leucine that was TCA-insoluble before dividing the culture, but that became TCA-soluble afterwards is referred to as the percentage of protein degraded.

The amount of protein degraded in the uninfected cultures varies as a function of the puromycin concentration used during the labelling period (Fig. 1). When no puromycin is added to the culture, 4% of the ¹⁴C-labelled proteins is degraded in 30 min; in the presence of 15 $\mu\text{g ml}^{-1}$, puromycin protein degradation in 30 min increases slightly to 5%; 50 $\mu\text{g ml}^{-1}$ puromycin produces a level of 14% protein degradation; and 100 $\mu\text{g ml}^{-1}$ puromycin 24%. The correlation between increasing concentrations of puromycin and increasing levels of protein degradation is expected; as the concentration of puromycin is increased there is an increase in the production of protein fragments that are turned over rapidly relative to normal proteins that are turned over slowly.

Although bacteriophage T4 infection has no detectable effect on the degradation of normal *E. coli* proteins (Fig. 1a), T4 infection drastically inhibits the degradation of ¹⁴C-PPF (Fig.

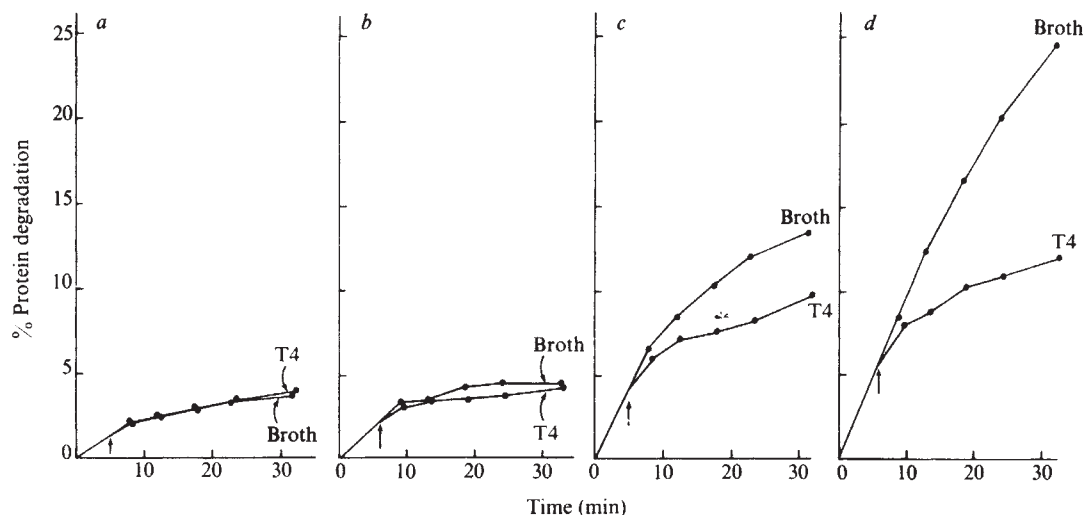
Table 1 PPF degradation after infection by T4, T5 or T7 phages

Culture*	% Protein degraded†
Uninfected	14
T4-infected	9
T5-infected	10
T7-infected	11

* *E. coli* B cultures growing at 37°C.

† The % of PPF degraded during the first 15 minutes after infection (or the comparable time interval for the uninfected control). Normalised data from two experiments are presented. Comparisons of PPF degradation at later time intervals are precluded by the short T7 latent period.

Fig. 1. Degradation of normal proteins and of puromycin protein fragments (PPF) in uninfected and in T4-infected *E. coli* CR63 cells. Puromycin ($0\text{--}100\text{ }\mu\text{g ml}^{-1}$) was added to mid-log phase CR63 cultures (5×10^8 cells per ml) in M9 medium¹⁶; 10 min later ^{14}C -leucine ($0.2\text{ }\mu\text{Ci ml}^{-1}$) was added. After a 5 min labelling period, the bacteria were collected on Millipore filters and washed twice to remove unincorporated puromycin and ^{14}C -leucine. The bacteria were then resuspended in fresh medium containing unlabelled leucine ($300\text{ }\mu\text{g ml}^{-1}$). The cultures were divided and one half of each was infected with wild type T4 D particles. Unless otherwise stated, the multiplicity of infection in this and in other experiments was 5 phage per cell, and the T4 particles used were wild type T4 D. Samples taken just before the cultures were divided (our 0-time samples) and at various times thereafter were added to cold TCA at a final TCA concentration of 5%. The TCA insoluble and soluble materials were separated from each other by low speed centrifugation. The amounts of ^{14}C -leucine in the TCA supernatants and pellets were determined by liquid scintillation counting techniques. The pellets were dissolved with Soluene (Packard) and then counted in BBOT scintillation fluid; aliquots of the supernatants were counted in Patterson Greene scintillation fluid⁷. The fraction of the ^{14}C -leucine that was TCA insoluble before dividing the culture, but that became TCA soluble afterwards is plotted as the percent of protein degradation. The arrows in this and in subsequent figures indicate the times when phages or other treatments were added to the bacterial cultures. a, No puromycin; b, $15\text{ }\mu\text{g ml}^{-1}$; c, $50\text{ }\mu\text{g ml}^{-1}$ and d, $100\text{ }\mu\text{g ml}^{-1}$ puromycin.



1b, c, d). Thus the degradation to acid soluble material of proteins made in the absence of puromycin is quantitatively similar in the T4-infected and in the uninfected cultures. As the concentration of puromycin increases, causing the production of progressively more PPF, the T4-induced inhibition of PPF degradation becomes apparent.

Degradation of normal and canavanine-containing proteins after T4 infection

To determine whether T4 infection inhibits the degradation of abnormal proteins other than puromycin protein fragments and to extend our observations of normal protein turnover after T4 infection, we performed the following double label experiment. Normal proteins in *E. coli* A-33, an arginine auxotroph, growing in the presence of arginine were labelled with ^3H -lysine; after a suitable labelling period, unincorporated ^3H -lysine and arginine were removed, and the growth medium was supplemented with unlabelled lysine and L-canavanine; 5 min later ^{14}C -leucine was added to label the abnormal canavanine-containing *E. coli* proteins (CCP). At the end of the labelling period, the cells were collected, washed free of unincorporated canavanine and ^{14}C -leucine, and then resuspended in medium containing arginine and unlabelled leucine. Degradation of the ^3H -normal proteins and ^{14}C -abnormal canavanine-containing proteins was measured in the T4-infected and in uninfected control cultures. The results of this experiment (Fig. 2) show no detectable difference in the degradation of normal ^3H -proteins between the T4-infected and uninfected cultures. Yet, in the same cells the degradation of abnormal ^{14}C -proteins is severely inhibited in the T4-infected culture compared with the uninfected control. We conclude that during T4 infection, the rapid degradation of abnormal proteins and protein fragments is inhibited, but the turnover of normal proteins is unaffected.

Inhibition of protein degradation after infection by various multiplicities of T4 particles

If the processes of T4 adsorption and injection were themselves directly responsible for altering the cell in such a way that the degradation of PPF and CCP were inhibited, then the extent of inhibition of degradation might be expected to increase with

increasing multiplicities of infection (MOI). If, on the other hand, T4 proteins synthesised after phage adsorption and injection were responsible for inhibition of degradation, little change in inhibition would be expected from different MOIs since only about four infecting T4 genomes per cell can be expressed in normal conditions.¹⁷ We therefore infected

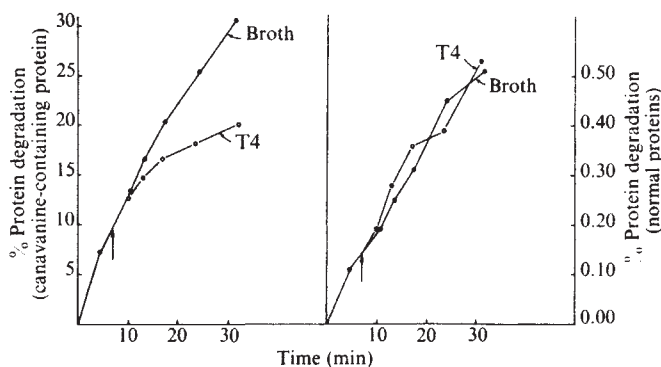


Fig. 2 The degradation of proteins in uninfected and in T4-infected *E. coli* A-33 cells containing both ^3H -normal proteins and ^{14}C -abnormal proteins. Normal proteins in *E. coli* A-33 bacteria growing in M9 medium supplemented with L-arginine ($50\text{ }\mu\text{g ml}^{-1}$) and L-tryptophan ($50\text{ }\mu\text{g ml}^{-1}$) were labelled with ^3H -lysine ($1\text{ }\mu\text{Ci ml}^{-1}$); the cells were then collected on filters, washed free of ^3H -lysine and of arginine, and resuspended in M9 medium lacking arginine, but supplemented with L-canavanine ($20\text{ }\mu\text{g ml}^{-1}$) and unlabelled lysine ($300\text{ }\mu\text{g ml}^{-1}$). After 10 min ^{14}C -leucine ($0.2\text{ }\mu\text{Ci ml}^{-1}$) was added to label the abnormal CCP; 5 min later the cells were again collected on filters, washed, and resuspended in fresh M9 with arginine ($100\text{ }\mu\text{g ml}^{-1}$) and unlabelled leucine ($300\text{ }\mu\text{g ml}^{-1}$). Thus, the *E. coli* A-33 culture contained ^3H -normal proteins and ^{14}C -CCP. One half of the culture was infected with T4 particles; degradation of ^3H -normal proteins and ^{14}C -CCP was assayed as described in Fig. 1. The low level of degradation of ^3H -normal proteins observed in this experiment is a consequence of the elapsed time between labelling normal proteins and the onset of our degradation assays. Those normal proteins that turn over rapidly had been degraded during the labelling of CCP and, therefore, had become TCA soluble by the start of our protein degradation measurements.

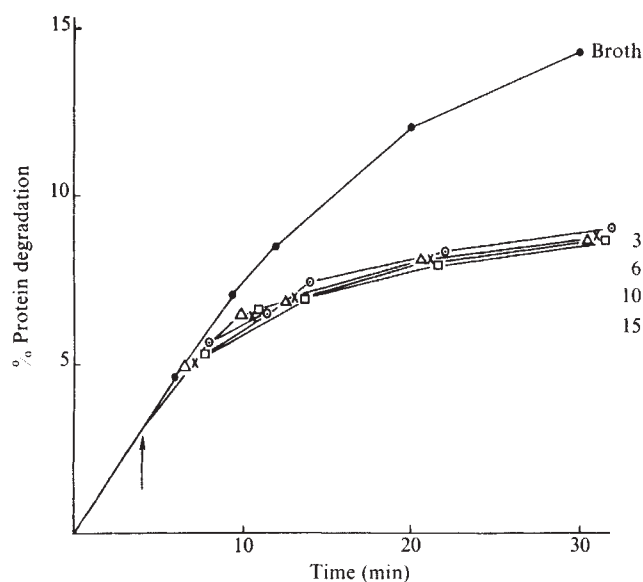


Fig. 3 The effect of various multiplicities of T4 particles on the degradation of *E. coli* puromycyl protein fragments. This experiment is similar to that in Fig. 1d, except that after exposure to puromycin ($100 \mu\text{g ml}^{-1}$) the CR63 culture was split into several aliquots which were then infected with the indicated multiplicities of T4 phages.

aliquots of a CR63 culture containing ^{14}C -PPF with various multiplicities of T4 particles. As can be seen in Fig. 3, the inhibition of PPF degradation following T4 infection is not multiplicity dependent. The same level of inhibition was observed with MOIs of 3, 6, 10, and 15 T4 particles per bacterium. Thus, it seems that the processes of T4 adsorption and injection are not directly responsible for the inhibition of PPF degradation.

Requirement for T4 protein synthesis to inhibit degradation of abnormal proteins

To find whether T4 protein synthesis is needed to block the degradation of *E. coli* protein fragments we used the antibiotic chloramphenicol (CAP) to inhibit protein synthesis at various times after T4 infection. Puromycyl protein fragments in *E. coli* CR63 were labelled with ^{14}C -leucine, as described above; the culture was divided and one half was infected with T4 particles. Our data (Fig. 4) show that T4 protein synthesis is required to inhibit the degradation of ^{14}C -PPF. If CAP is added and T4 protein synthesis is blocked 1 min after the addition of T4 particles, the T4 infected and uninfected cultures show essentially no difference in PPF degradation. CAP does not interfere with T4 adsorption or injection; over 95% of the cells in the 1-min CAP sample had lost their colony-forming ability because of T4 infection. If CAP is added to samples at progressively later times after T4 infection, the samples show progressively more inhibition of PPF degradation. The degradation of PPF when CAP is added 7 min after infection is nearly the same as in the T4-infected sample with no CAP. The addition of CAP 10 min or longer after infection does not affect the inhibition of PPF degradation (data not shown). Therefore, it seems that T4 proteins synthesised early after infection inhibit the degradation of puromycyl peptides. Further, the continued synthesis of these proteins is not necessary to sustain the inhibition for the duration of the latent period.

Maturation-defective T4 mutants¹⁵ have also been used to demonstrate that T4 early proteins are sufficient to inhibit PPF degradation. The proteins encoded by T4 genes 33 and 55 are required continuously for the synthesis of late T4 mRNA; thus, in T4 33⁻ or 55⁻ infections only early proteins are produced.

Our results (Fig. 5) show that in T4 33⁻ or 55⁻ infected *E. coli* cells PPF degradation is inhibited to the same extent as in the control T4 infection. We conclude that inhibition of PPF degradation is an early T4 function.

The extent of T4-induced inhibition

It is difficult to assess with certainty the extent to which T4 inhibits CCP and PPF degradation. The continued solubilisation of labelled proteins that we have observed after T4 infection is at least in part due to (1) the time required for T4 particles to adsorb to bacteria, to inject their DNA, and to effect maximum

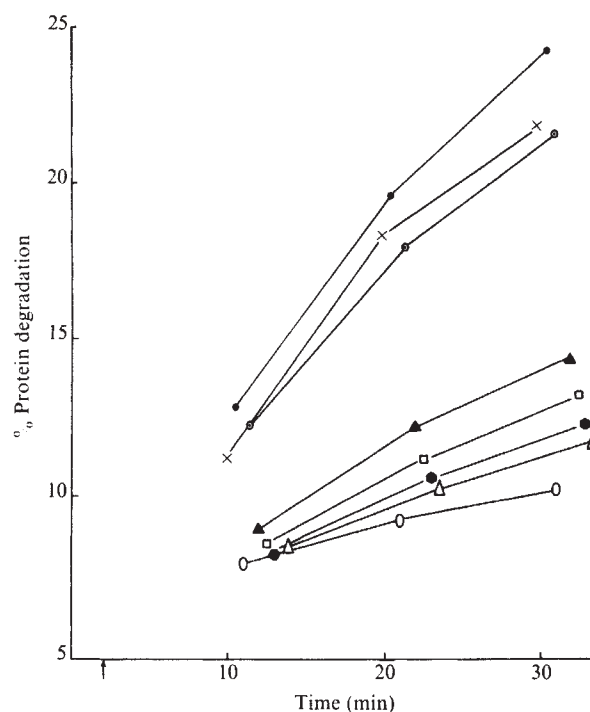


Fig. 4 The requirement for T4 protein synthesis to inhibit degradation of puromycyl protein fragments. This experiment is similar to that described in Fig. 1d, except protein synthesis was blocked at 1 min (○), 4 min (▲), 5 min (□), 6 min (●), and 7 min (△) after T4 infection and immediately after dividing the CR63 culture in one of the uninfected controls (●) by addition of chloramphenicol (CAP) at a final concentration of $100 \mu\text{g ml}^{-1}$. x, Uninfected control; 0, infected control.

inhibition, (2) the uninterrupted turnover of those normal proteins synthesised along with PPF or CCP during the labelling period, and (3) continued cleavages by peptidases of CCP and PPF molecules that were substantially degraded before T4-induced inhibition became effective. As treatment of cells with organic solvents such as toluene virtually eliminates the degradation of β -galactosidase amber fragments (D. Zipser, personal communication), we were interested to compare the inhibition of PPF degradation observed after treatment with organic solvents to that observed after T4 infection (Fig. 6). The organic solvent used was chloroform, which blocks degradation as rapidly as any treatment we have tested. Our results indicate that addition of chloroform to *E. coli* cells inhibits PPF turnover only slightly more rapidly than does addition of T4 particles (MOI = 5). This difference in kinetics of inhibition is probably due to the longer time required by T4 to maximally block PPF degradation. The turnover rates of PPF in the T4 infected and chloroform treated samples are nearly the same by 10 min after infection.

Inhibition of protein degradation by T5 and T7 phages

Inhibition of degradation of abnormal proteins and protein fragments following virus infection is not peculiar to T4 phages. Other phages, unrelated to T4, such as T5 and T7, also inhibit PPF degradation (Table 1). Neither T5 nor T7 infection, however, seems to inhibit PPF breakdown to as great an extent as does T4 infection.

Implications

We have found that infection by T4, T5, or T7 bacteriophages results in the inhibition of degradation of abnormal *E. coli* proteins (CCP) and puromycinyl protein fragments (PPF). Normal *E. coli* proteins, however, continue to turn over at their usual slow rates despite T4 infection. The differential effects of T4 on the turnover rates of normal and of abnormal proteins are not explained by the model of one *E. coli* proteolytic system in which the $V_{\max}$ of a crucial degradative enzyme is reduced after phage infection so that at high substrate concentration (for example, in the presence of PPF or CCP) the reduced $V_{\max}$ becomes rate limiting, whereas at low substrate concentrations (in the absence of PPF or CCP) substrate is limiting. On the contrary, our experiments (Fig. 2) show that in the same T4-

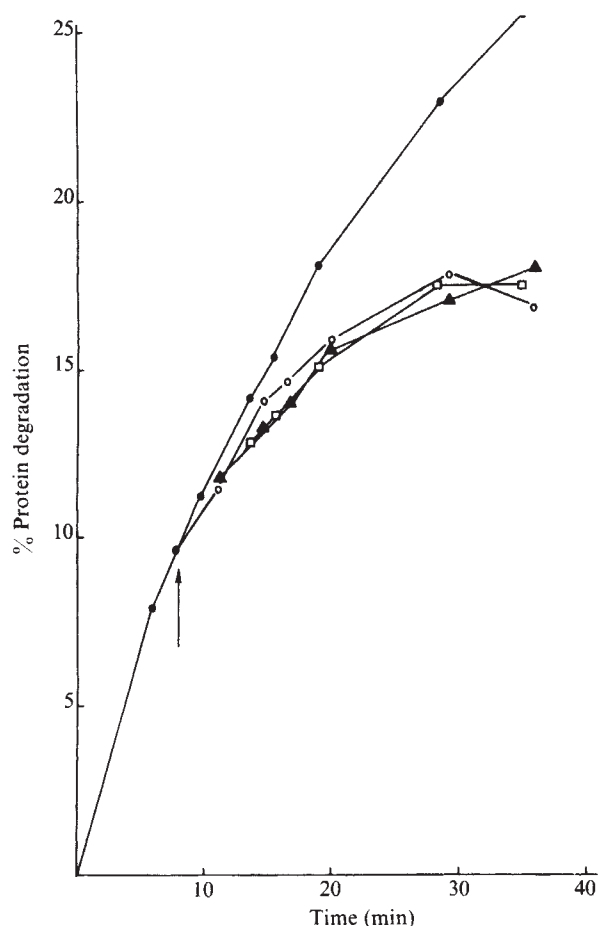


Fig. 5 Inhibition of PPF degradation after infection by T4 maturation defective mutants. This experiment is similar to that described for Fig. 1d with the following exceptions: an *E. coli* B (*Lon*⁺) culture containing ¹⁴C-PPF was divided into four aliquots, one of which was infected with T4am N134 (gene 33⁻) (○), another with T4am BL292 (gene 55⁻) (▲), the third with wild type T4 particles (□), and the fourth was treated only with broth (●). Degradation of ¹⁴C-PPF in each aliquot was assayed as described in Fig. 1.

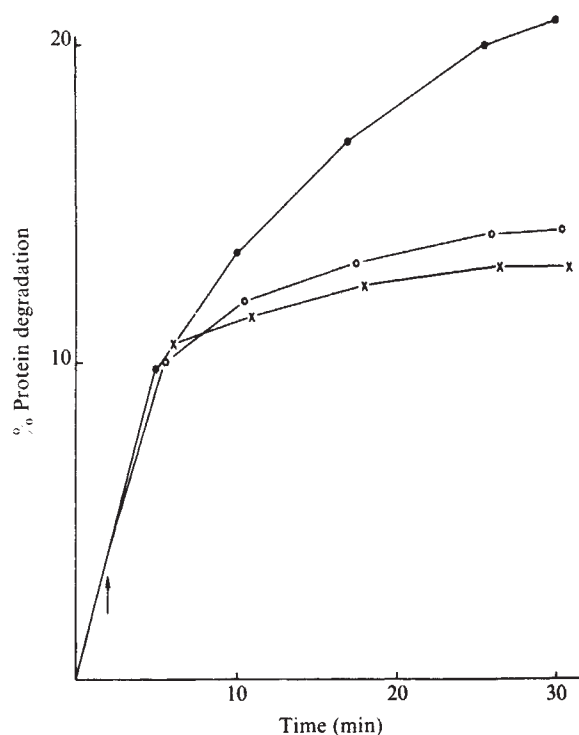


Fig. 6 The effect of chloroform on degradation of PPF. This experiment is similar to that described for Fig. 1d, except that several drops of chloroform were added to one aliquot of the ¹⁴C-PPF-containing *E. coli* CR63 culture. ●, Treatment with broth only; ○, T4 infection only; ×, chloroform added.

infected cells at the same times that CCP (or PPF) degradation is inhibited, the degradation of normal proteins is unaffected. Thus, high substrate concentrations are not related to the differential effects of T4 infection on the degradation of normal and of abnormal proteins. Our results suggest that, at least in part, the *E. coli* pathway for the degradation of abnormal proteins is distinct from the pathways for the degradation of normal cellular proteins. Therefore, the turnover of most normal *E. coli* proteins is not a consequence of changes in these proteins which renders them susceptible to degradation as abnormal proteins.

Our observations explain why nonsense fragments of most T4 proteins are readily found in extracts of infected cells⁶, whereas *E. coli* nonsense fragments are usually not found due to their rapid degradation in uninfected cells. This difference in turnover rates is not the result of some intrinsic distinction between T4 and *E. coli* proteins, but rather of the inhibition of CCP and PPF degradation following T4 infection.

Inhibition of CCP and PPF degradation requires T4 protein synthesis. Adsorption of T4 particles to the *E. coli* surface and injection of T4 DNA are not by themselves sufficient to alter protein turnover. The results of experiments using chloramphenicol to block T4 protein synthesis at various times after infection suggest that an early T4 function is responsible for inhibition of CCP and PPF degradation. This suggestion is consistent with our observations following infections by T4 maturation defective mutants. In T4 33⁻ or T4 55⁻ infection, early proteins, but no late proteins, are made; nevertheless, degradation of CCP and PPF is inhibited to the same extent as in wild type T4 infection. We conclude, therefore, that early T4 protein(s) shut off the *E. coli* system for degradation of abnormal proteins and protein fragments.

The fact that several unrelated phages are able to inhibit CCP and PPF degradation in their host cells suggests that this ability may have some evolutionary significance for the survival of these phages. Perhaps specific proteins of various phages and

other viruses would be degraded as though they were abnormal proteins by the host cell if its degradation system were not inhibited.

This work was supported by grant PCM 77-08112 from the NSF to L.D.S., by a Busch Predoctoral Fellowship to K.T., and by a PHS Biomedical Research Support grant to A.S.J.

Received 31 May; accepted 7 August 1978.

1. Goldberg, A. L. & Dice, J. F. A. *Rev. Biochem.* **43**, 835-869 (1974).
2. Goldberg, A. L. & St John, A. C. A. *Rev. Biochem.* **45**, 747-803 (1976).

3. Goldschmidt, R. *Nature* **228**, 1151-1154 (1970).
4. Lin, S. & Zubin, I. *J. biol. Chem.* **247**, 2205-2211 (1972).
5. Zipser, D. & Bhausaar, P. *J. Bact.* **127**, 1538-1542 (1976).
6. Celis, J. E., Smith, J. D. & Brenner, S. *Nature new Biol.* **241**, 130-132 (1973).
7. Goldberg, A. L. *Proc. natn. Acad. Sci. U.S.A.* **69**, 422-426 (1972).
8. Kowit, J. D. & Goldberg, A. L. *J. biol. Chem.* **252**, 8350-8357 (1977).
9. Prouty, W. F. & Goldberg, A. L. *Nature new Biol.* **240**, 147-150 (1972).
10. Bukhari, A. I. & Zipser, D. *Nature new Biol.* **243**, 238-241 (1973).
11. Gottesman, S. & Zipser, D. *J. Bact.* **133**, 844-851 (1978).
12. Schachtele, C. F., Anderson, D. L. & Rogers, P. *J. molec. Biol.* **33**, 861-872 (1968).
13. Prouty, W. F., Karnovsky, M. J. & Goldberg, A. L. *J. biol. Chem.* **250**, 1112-1122 (1975).
14. Simon, L. D. *Proc. natn. Acad. Sci. U.S.A.* **69**, 907-911 (1972).
15. Bolle, A., Epstein, R. H., Salser, W. & Geiduschek, E. P. *J. molec. Biol.* **33**, 339-362 (1968).
16. Adams, M. *Bacteriophages* (Wiley, New York, 1959).
17. Snustad, D. P. *Genetics* **54**, 937-954 (1976).

letters to nature

An ultraviolet subdwarf companion to HD17576

ULTRAVIOLET objects detected by the S2/68 UV sky-survey telescope¹ aboard the TD-1 satellite consist primarily of known early-type stars. In addition, there is a substantial subgroup of unidentified and/or untyped ultraviolet stars including some new hot subdwarfs^{2,3} and a small subgroup of known late-type objects including the 20 stars of types F and G listed in the Ultraviolet Bright-Star Spectrophotometric Catalogue⁴. Selection of those objects whose ultraviolet colours suggest that they are earlier than an unreddened B5 V star ought to eliminate the latter category entirely, but several apparently late-type anomalies persist. If these objects cannot be discarded on grounds of blending, misidentification, inaccurate ultraviolet colours (objects with low flux or few orbital passes), or erroneous prior spectral types taken from the literature, then one is confronted with the possibility of genuine duplicity. UV0246-37, taken from the UCL Catalogue of Ultraviolet Objects soon to be published, is just such an object: a binary system consisting of a previously unrecognised hot ultraviolet star and a cooler G-type primary which wholly dominates in the visible region.

Six unblended spectral scans of UV0246-37 were obtained by the S2/68 experiment in 1972-3, two of which must be rejected: in one pass the target lay too far off-centre to yield reliable data, and in the other the data were contaminated by the enhanced radiation background of the South Atlantic Anomaly⁵. Photon pulse counts were obtained in three 400-Å spectrometer channels centred on 1,550 Å (A2), 1,950 Å (A3), and 2,350 Å (A4), and in a photometric channel with FWHM 310 Å centred on 2,740 Å. These counts were converted to flux densities through the absolute calibration determined by Humphries *et al.*⁶ The mean ultraviolet spectrum from the four passes is plotted in Fig. 1, excluding the data at 1,750 Å and 2,150 Å where vignetting was severe. The steep blueward rise is reflected in the mean difference between the magnitudes derived from pulse counts in the A1 and A2 channels: $m_{A2} - m_{A1} = -2.00$.

The position of UV0246-37 coincides with that of HD17576 (=DM-37° 1050 = SAO193923), a star of $m_{pg} = 8.1$ classified as G0 on two HD plates⁷. To verify the spectral classification and check for evidence of duplicity, uvby and β photometry was conducted on the 1-m telescope at the South African Astronomical Observatory in July 1977; and spectra at reciprocal dispersion 30 Å mm⁻¹ (second-order blue) were obtained on backed Ila0 plates with the f/1.4 EMI image-tube spectrograph at Cassegrain focus on the 1.9-m telescope in September 1977.

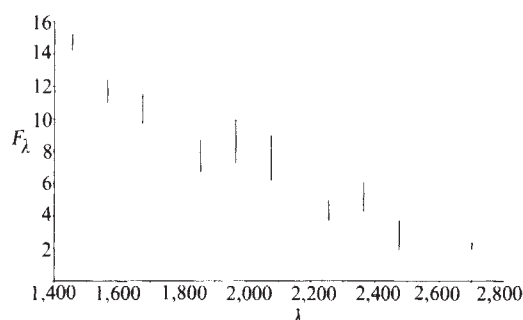
The luminosity class of the G0 star—as predicted among other factors on the strength of the G band, the dominance of H δ over the Fe I and Sr II lines shortward of it, and measures of luminosity-sensitive line ratios⁸—must be V or IV-V. The only spectroscopic trait indicative of the presence of a hot companion is the ratio of the Ca II lines (K:H $\approx$ 3:4), long held to be a reliable diagnostic of a composite spectrum⁹ when the 'cool' primary is later than type F and the secondary earlier than type A.

The composite nature of HD17576 is strikingly confirmed by the uvby photometry (for a description of the uvby system, including the passbands of the four filters, see ref. 10):

$$\begin{aligned} V &= 7.82 \pm 0.02; u - b = 1.152 \pm 0.006 \\ b - y &= 0.382 \pm 0.006; m_1 = 0.128 \pm 0.010 \\ c_1 &= 0.113 \pm 0.010; \beta = 2.584 \pm 0.007 \end{aligned}$$

(weighted means of three observations). Since m_1 is low and c_1 very low indeed compared with the approximate mean values (for G0 V) of 0.18 and 0.34 respectively (E. Oblak, personal communication and ref. 11), it stands out conspicuously in a plot of the reddening-insensitive indices [c_1] against [m_1] (Fig. 2). On the other hand, $(b - y) = 0.382$ is not anomalous and though not a reliable luminosity discriminant further supports the G0 V classification of HD17576 since $(b - y)$ remains virtually unaffected by any faint hot companion. [$u - b$] = [c_1] + 2[m_1] = 0.369 is certainly bluer than normal (~ 0.76); had it not been

Fig. 1 Mean ultraviolet spectrum of UV0246-37 from the S2/68 experiment: nine unvignetted points in the wavelength region covered by the spectrophotometer channels (A2, A3, and A4) and the photometric datum at 2,740 Å (A1 channel). Ordinate units are 10^{-12} erg cm⁻² s⁻¹ Å⁻¹; abscissa units, Å. Error bars indicate the 1 σ deviation over four of the six recorded orbital passes (see text).



contaminated, it could well have served as a luminosity discriminant¹¹. The H β measurement is slightly low (Crawford¹² tabulates 2.606 as the norm for G0) but among the G stars the β index loses most of its usefulness as a luminosity or temperature discriminant through the crowding of absorption lines.

The presence of the ultraviolet star in a binary system with a normal late-type primary enables us to infer the absolute luminosity of the hot secondary. The standard luminosity calibration gives $M_v = 4.4$ for a main-sequence G0 star^{13,14}; with $m_v = 7.82$, the distance turns out to be only ~ 50 pc.

On the absolute flux calibration scale determined for Vega by Hayes and Latham¹⁵, the apparent magnitude of UV0246–37 at 2,740 Å is $m_{A1} = 8.07$. To derive m_v for the hot secondary, the UV flux curve can be extrapolated—a possibly reprehensible procedure in view of the 0.2 or 0.3 mag uncertainty incurred by the Balmer jump—to give $(m_v - m_{A1}) = 2.0$. Alternatively, we can use the fact that the flux distribution of UV0853+01, an early-type field star unaffected by the presence of a cool companion, mimics that of UV0246–37 to infer the same 2.0 mag difference. Moreover, if the mean value $\overline{m_v - m_{A1}}$ is calculated for all hot subdwarfs (J. R. Giddings, unpublished data) with ultraviolet colours $(m_{A2} - m_{A1})$ and $(m_{A2} - m_{A4})$ deviating no more than ± 0.2 from those for UV0246–37, we arrive once more at the same result.

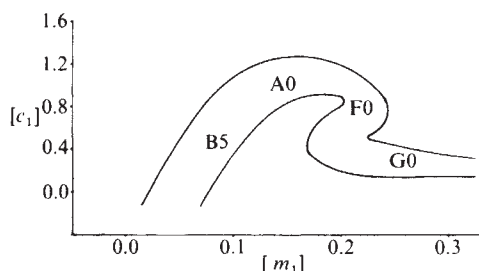


Fig. 2 Boundaries of the main sequence in the $[c_1]$ – $[m_1]$ plane, with normal positions for several spectral types shown. In standard notation¹⁰, $[c_1] \equiv c_1 - 0.20(b - y)$ and $[m_1] \equiv m_1 + 0.18(b - y)$. The diamond represents HD17576, whose anomalous position reflects its composite nature. The main-sequence band is based on the observational material of Oblak *et al.*¹¹.

Thus $m_v \approx 10.1$ for the secondary and its absolute magnitude follows: $M_v = 6.7$. If the determination of MK luminosity were in error by half a class interval, a not uncommon discrepancy¹⁶, then the quoted value could be 0.6 mag too faint. In any case it is at once evident that the secondary is markedly subluminescent for an early-type star albeit more luminous than any white dwarf; it is indeed a hot subdwarf akin to such objects in the UCL Catalogue as UV1758+36 (ref. 2), UV0921+36 (BD +37° 1977) (refs 17–19), and UV2158–02 (BD –3°5357) (ref. 20) but far more subluminescent. Nevertheless, it is not without precedent: four objects in the list of subdwarf O stars by Greenstein and Sargent²¹ (HZ 38, HZ 3, Feige 26, and BD +25°4655) have $M_v \geq 6.0$. Luyten²² considers the kinematic data for O and B subdwarfs scarcely distinguishable and accordingly treats them as a single group for which, on the basis of his proper-motion survey of faint blue stars, he finds $M_v = 4.1$. Evidently UV0246–37 has an exceptionally low luminosity, even for a subdwarf, indeed approaching that of the most luminous white dwarfs. However, the values of M_v for the hottest white dwarfs are ill determined: for HZ 43, assigned $M_v = 7.1$ by Greenstein and Sargent²¹, a careful study by Margon *et al.*²³ finds $M_v = 9.1 \pm 0.5$. Subdwarfs are in any case readily distinguished from white dwarfs on photometric and spectroscopic criteria (for example, ref. 24) except in a special case like the present one where the required observational data cannot be simply ascertained.

At this stage, estimation of the physical parameters of UV0246–37 is an invidious task inasmuch as the errors in the flux measurement are substantial and T_{eff} estimates from a black-body model are notoriously sensitive to the slope (for example, ref. 23). A crude black-body fit yields a temperature of $4.2^{+0.5}_{-0.7} \times 10^4$ K. It may be noted that HZ 38, an O subdwarf with the same luminosity²¹ as the one at hand, has $T_{\text{eff}} = 36,000$ K and $\log g = 7.0$.

The significance of the discovery reported here is twofold. Carnochan *et al.*¹⁷ have mustered evidence from the faint (limiting $m_v \sim 10$), unclassified ultraviolet objects detected by S2/68 that there exists a substantial population of subluminescent hot stars in a region of the HR diagram previously considered sparsely populated. Their space density and scale height have been calculated and will be published elsewhere. UV0246–37 is manifestly drawn from this population.

Second, there is a growing body of recognised cool stars with hot secondaries^{20,25,26}, few of which represent such extreme cases of spectral segregation as UV0246–37, the primary being undetectable in the ultraviolet region and the presence of the secondary nearly indiscernible in the visible. The S2/68 survey yielded several other late-type ultraviolet anomalies of which ground-based spectroscopy and photometry are being carried out (ref. 27 and J. P. Swings and M. Berger, personal communications), and it is anticipated that other cool stars will prove to have unsuspected hot companions. These objects are prime candidates for investigation by the International Ultraviolet Explorer satellite since their fortuitous occurrence in binary systems will help to elucidate the nature of the hot subdwarfs.

We thank D. J. Carnochan and J. R. Giddings for discussions and assistance in data reduction. The hospitality of the South African Astronomical Observatory and the Royal Greenwich Observatory is acknowledged. This work was supported by the SRC and by a Royal Society Radcliffe Lectureship to J.D.

JON DARIUS
PATRICIA A. WHITELOCK*

Department of Physics and Astronomy,
University College London,
London WC1, UK

Received 6 June; accepted 28 July 1978.

* Present address: South African Astronomical Observatory, PO Box 9, Observatory, Cape 7935, Republic of South Africa.

1. Bokserberg, A. *et al. Mon. Not. R. astr. Soc.* **163**, 291–322 (1973).
2. Giddings, J. R. & Dworetzky, M. M. *Mon. Not. R. astr. Soc.* **183**, 265–269 (1978).
3. Giddings, J. R. *et al.* (in preparation).
4. Jamar, C. *et al. ESA Scient. Rep.* No. 27 (1976).
5. Vernov, S. N., Gorchakov, E. V., Shavrin, P. I. & Sharvina, K. N. *Space Sci. Rev.* **7**, 490–533 (1967).
6. Humphries, C. M., Jamar, C., Malaise, D. & Wroe, H. *Astr. Astrophys.* **49**, 389–406 (1976).
7. Cannon, A. J. & Pickering, E. C. *Harvard Ann.* **91**, 191 (1918).
8. Buscombe, W. & Dickens, C. R. *Mon. Not. R. astr. Soc.* **128**, 499–507 (1964).
9. Hynek, J. A. *Contr. Perkins Obs.* **1**, 185–256 (1938).
10. Golay, M. *Introduction to Astronomical Photometry* (Reidel, Dordrecht, 1974).
11. Oblak, E., Considère, S. & Charetton, M. *Astr. Astrophys. Suppl.* **24**, 69–88 (1976).
12. Crawford, D. L. *Dudley Obs. Rep.* No. 9, 17–29 (1975).
13. Allen, C. W. *Astrophysical Quantities*, 3rd edn (Athlone, London, 1973).
14. Gliese, W. *Veröff. astr. Rechen-Inst. Heidelberg* No. 24 (Braun, Karlsruhe, 1971).
15. Hayes, D. S. & Latham, D. W. *Astrophys. J.* **197**, 593–601 (1975).
16. Jaschek, C. & Jaschek, M. *IAU Symp.* No. 24, 6–12 (1966).
17. Carnochan, D. J., Dworetzky, M. M., Todd, J. J., Willis, A. J. & Wilson, R. *Phil. Trans. R. Soc. A279*, 479–485 (1975).
18. Wolff, S. C., Pilachowski, C. A. & Wolstencroft, R. D. *Astrophys. J. Lett.* **194**, L83–L85 (1974).
19. Viotti, R., Altamore, A. & Baratta, G. B. *Astr. Astrophys.* **61**, 133–136 (1976).
20. Dworetzky, M. M., Lanning, H. H., Etzel, P. B. & Patenaude, D. J. *Mon. Not. R. astr. Soc.* **181**, 13P–18P (1977); *IAU Inf. Bull. Variable Stars* No. 1284 (1977).
21. Greenstein, J. L. & Sargent, A. I. *Astrophys. J. Suppl.* **28**, 157–210 (1974).
22. Luyten, W. J. *Proper Motion Survey with the 48-inch Schmidt Telescope*, No. 45 (University of Minnesota, Minneapolis, 1976).
23. Margon, B. *et al. Astrophys. J.* **209**, 525–535 (1976).
24. Wickramasinghe, D. T., Bessell, M. S. & Cottrell, P. L. *Astrophys. J. Lett.* **217**, L65–L68 (1977).
25. Parsons, S. B., Wray, J. D., Kondo, Y., Henize, K. G. & Benedict, G. F. *Astrophys. J.* **203**, 435–437 (1976).
26. Kohoutek, L. & Laustsen, S. *Astr. Astrophys.* **61**, 761–763 (1977).
27. Darius, J. *et al.* (in preparation).

The origin of lunar palaeomagnetism

THE new determination of magnetic field anomalies over part of the Moon's surface from Apollo 15 and 16 subsatellite magnetometer data has recently been interpreted¹ in terms of magnetised areas of the lunar surface. I show here that, from this analysis, palaeomagnetic pole positions can be calculated and that these are so clustered on the lunar surface that there is evidence against meteoritic or cometary processes as explanations of the remanent magnetisation of the Apollo rocks. I conclude, therefore, that the Moon had a magnetic field in its early history produced by dynamo processes in a fluid electrically conducting core.

Since the discovery of remanent magnetisation of Apollo 11 lavas², we have attempted in a series of papers to interpret lunar palaeomagnetism as observed in the remanent magnetisation of the returned Apollo crystalline rocks and breccia and in widespread crustal magnetic anomalies mapped by magnetometers on Explorer 35³ and the Apollo 15 and 16 subsatellites⁴, and by the reflected electron technique⁵. Surface traverses of Apollo 14 and 16⁶ also detected magnetic anomalies. Our interpretation supposed that the Moon once had a magnetic field in which these rocks acquired, at the time of their cooling from above the Curie point of their iron grains, a stable thermoremanent magnetisation. This field existed at least from 4,000 to 3,200 Myr (the ages of the rocks studied), decreased from 1 G to 5,000 nT (ref. 7) over this interval and has now disappeared. Much scepticism has been expressed⁸ about this hypothesis of an ancient internal field, presumably because a number of consequences follow which had not been envisaged by the majority of lunar scientists—the existence of a sizeable iron core in which a dynamo process can take place⁹, the melting of the entire Moon¹⁰ in its earliest history (so that the iron core formed) and thermal convection in the interior.

Consequently there have been several suggestions as to how the magnetisation of the lunar rocks might have been acquired by an external field¹⁰ (or no field at all, although as T. Gold points out, some far-reaching modification in physics, for example, the law of parity, would seem to be necessary if this explanation were to be sustained). These ideas appeal to processes with which we are unfamiliar in terrestrial studies, and they are therefore difficult to test. For example, the ubiquity of large meteorite impacts on the early lunar crust has suggested the possibility of impact magnetisation^{12,13}. Also it has been pointed out that cometary impacts are statistically likely to have been quite frequent on the Moon and again the physical processes during such events may have favoured magnetisation¹⁴. What must be postulated is that the acquisition of a stable remanent magnetisation is greatly facilitated during such large impacts either by the effect of the transient stress field on the

iron grains (although the magnetisation direction might be biased by the stress) or through the shock wave raising the rock temperatures instantaneously to above the blocking temperature of the iron grains. Thus the magnetising field on such a hypothesis could be small and transient: it may arise from surges of electric current in the plasma during impact, 'pseudo-lightning strikes', or be the field of the solar wind, which does not remain constant for long. We are concerned with the early history of the Solar System when the solar wind field may have been much more intense than at present. The average solar wind field now lies in the ecliptic, therefore a magnetisation of the Moon along its rotation axis is implausible if the magnetisation was acquired very quickly as must be the case in an impact magnetisation process¹⁵. On the other hand, because of the Moon's rotation, a magnetisation in the ecliptic or the lunar equatorial plane is unlikely unless the magnetisation was acquired in a time much shorter than a month.

Very important interpretations of the lunar magnetic anomalies, as seen by the Apollo 15 and 16 subsatellites, which bear on these critical questions, have been made by Hood, Coleman and Russell¹. They show that these magnetic anomalies at altitudes of 100 km can be represented by 35 dipoles at the surface of the Moon and they calculate the strength and direction of these dipoles. Strangway¹⁶, and Dermott and Runcorn (unpublished) found that the largest magnetic anomalies of the lunar far side demand east-west directed dipoles for their explanation. Hood *et al.* show that the dipoles are not randomly directed in space but lie nearly in the equatorial plane of the Moon. However, they show that their directions, projected in this plane, are scattered, apparently randomly, through 180°. They conclude that this random feature is the explanation of the null result found in the determination of the present dipole moment of the Moon¹⁷. This result I had explained¹⁸ by showing that a spherical shell permanently magnetised by a field of internal origin, so that the magnetisation is parallel and proportional to the field, which later disappears, produces no external field. I had argued that such magnetic anomalies would arise from the large craters releasing lines of magnetic force from the magnetised lunar shell. Thus the anomalies should be well modelled by a dipole in the centre of the crater but it has been often found in such attempts (ref. 1 and Dermott and Runcorn, unpublished) that the dipoles are not so positioned, although the largest magnetic anomalies are associated with the large craters around Van de Graaf on the far side. It therefore seems likely that the magnetised areas may sometimes result from the cooling of layers of breccia thrown out by meteoritic impacts, as suggested by Strangway *et al.*¹⁶.

The two features of the distribution of dipole directions, that they are near to the lunar equator and in that plane are randomly directed, seem to suggest that a field is responsible which is directed in this plane and this suggests the solar wind field which, though fluctuating, is on average in the plane of the ecliptic, to which the lunar equator is inclined by only 1°. Cometary or

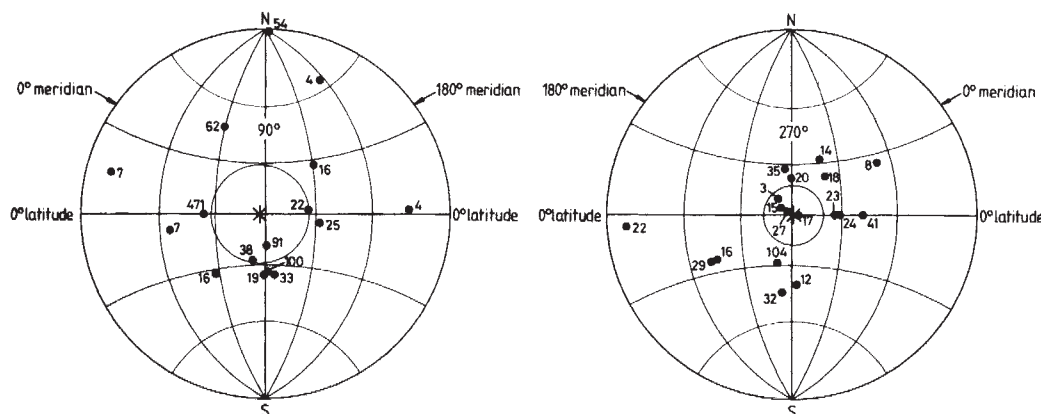


Fig. 1 Distribution of pole positions from lunar magnetic anomaly dipoles (given in Table 1 of ref. 1). Stars are mean directions and circles the 95% circles of confidence. Stereographic projection of eastern (a) and western (b) hemispheres of Moon. The numbers are the dipole strengths per unit area given in G cm. Left, eastern hemisphere. Right, western hemisphere.

Table 1 Number of poles within 30° sectors of longitude compared with number expected on the theory that the dipoles are randomly orientated in the lunar equatorial plane

Longitude sectors	No. of poles all included		No. of poles magnitudes <10 G cm excluded	
	Observed	Theory	Observed	Theory
Eastern Hemisphere				
0–30°, 150°–180°	2	3.1	0	2.3
30°–60°, 120°–150°	6	3.7	4	2.7
60°–90°, 90°–120°	7	6.0	7	5.9
Western Hemisphere				
180°–210°, 330°–360°	1	3.8	1	3.6
210°–240°, 300°–330°	4	4.5	4	4.2
240°–270°, 270°–300°	13	9.6	12	9.1

$\chi^2 = 5.30$ (all poles included) and 5.95 (excluding <10 G cm). Confidence limits for rejection of theory, 75% and 80% respectively.

meteoritic impact hypotheses seem to be supported. However, Table 1 of Hood *et al.*¹ gives the locations on the Moon of these dipoles and their angles of orientation to the vertical and the meridian are listed. I have calculated the corresponding pole positions, as has proved fruitful in palaeomagnetism, that is, the latitudes and longitudes of the axes of those dipoles at the centre of the Moon which have fields at the various locations coincident with the palaeofield directions¹⁹ found there. These poles are plotted in Fig. 1 and are seen not to be random. A minority of directions are opposed to the most frequent grouping: this is to be expected as a dipole may either arise from a magnetised body of rock or from a magnetised part of the lunar crust in which there is a demagnetised volume, a physical cavity formed by impact or a volume of rock demagnetised by impact. Moreover, reversals of polarity of a field produced by a core dynamo process is to be expected. The scatter of the pole positions is considerable but this is to be expected as many of the anomalies, so remarkably analysed and modelled by Hood *et al.* are much less than 1 nT, and their treatment of the magnetometer data has required the removal of the much larger field of the Earth's magnetotail, which is not very constant. In fact the nonrandom distribution of the pole positions demonstrates that this difficult process has been successfully accomplished by Hood *et al.* and suggests that an attempt to separate the magnetic anomalies from other regions of the Moon, at present obscured by the solar wind field, might be worth attempting.

Thus the demonstration that the vector means of the two groups of pole positions, shown with their 95% circles of confidence in Fig. 1, are at 180°, seems at first sight to be strong evidence that the magnetisation of the lunar crust does not arise from meteoritic or cometary impacts, for the palaeomagnetic directions impressed on the lunar crust from such processes must, because of the lunar rotation, be instantaneous directions from the solar wind magnetic field. Unless all the impacts occurred at the same time of the month, the palaeomagnetic directions over the Moon as a whole would not have an axis fixed in one region of the Moon. However, the fact that the mean poles are along the E–W equatorial axis is odd. Consider a set of dipoles at a site (0° N, 180° E) lying in the equatorial plane. If the difference in azimuths of two neighbouring dipoles is $d\theta$ then their poles lie on the equator at distance $d\theta$ apart where θ is the angular distance from the site

$$d\theta = dI(1 + 3 \cos^2 \theta)/2$$

Thus if these dipoles at one site are distributed randomly through 180°, then their poles are clustered four times more densely about an E–W equatorial axis than about the N–S axis. If random N–S components are added, this distribution would resemble that of Fig. 1. However, I show now that the conclusion that the dipoles are random is wrong, although clearly the dipole

directions will be substantially affected by random components inherent in the magnetic survey observations and analyses. It is significant that four out of six of those poles which lie far away from the main group have dipole moment per unit area magnitudes less than 10 G cm, that is corresponding to fields less than 0.1 nT at the subsatellite orbit.

The sites of the dipoles are in fact distributed over a strip 45° on either side of the mean site at 10° S 180° E. If random dipoles were uniformly distributed over the equator between 135° E and 215° E then the pole density distribution would be proportional to $\tan^{-1} \sec 2\theta$, with a maximum at 90° E and 270° E, three times that of the minimum at 0° and 180°.

On the above hypothesis the proportion of poles within the 30° ranges of longitudes is shown in Table 1 and is compared with the actual number. It is clear that χ^2 test excludes at the 80% confidence level the hypothesis that the directions are wholly random: an east–west component is also present.

Yet there is an external field to which the Moon is exposed during 2–3 days around full Moon and which is directed along the Earth–Moon line. This is the Earth's magnetotail: the Moon traverses both its north and south lobes, in which the magnetic field is directed towards and away from the Sun respectively, in the course of the year. Thus if impacts had imparted to crustal regions magnetisations directed along this field, a plot resembling Fig. 1 would be found. However, the Moon spends 10 times as long in the solar wind as in the magnetotail, and so unless impacts occurred always at full Moon the crustal magnetisations should be nearly randomly distributed in the ecliptic. This hypothesis would, if 90° of rotation of the Moon about its axis since magnetisation had occurred, explain the results of Table 1, had the magnetised area been small. But for two areas of longitude difference $d\lambda$ within the strip, imprinted by the magnetotail field, their calculated pole positions would be $d\theta$ apart where $d\theta = 3d\lambda(2 + \tan^2 \lambda)/(4 + \tan^2 \lambda)$. This pole distribution is spread nearly uniformly over about 70° on either side of the mean pole. Thus the additional hypothesis of magnetisation by the magnetotail field, as well as the solar wind field, would broaden the pole distribution of the theory columns of Table 1, rather than reduce it as required by the observations.

So in spite of random scatter and the restricted area covered by the anomalies, it is concluded that the ancient pole position was not near the present axis of rotation of the Moon. Yet, in spite of the lunar core being an order of magnitude smaller than the Earth's core and the angular rotation less by a factor of 30, the reason originally advanced to explain the fact that the average geomagnetic field is directed along the axis of the Earth's rotation, that is the Coriolis force is many orders of magnitude greater than any other term in the hydrodynamic equation²⁰, must also apply to the Moon. Therefore, I conclude that lunar palaeomagnetism shows that polar wandering occurred in the Moon's early history, that is, the Moon rotated relative to its axis of rotation. It was shown previously that polar wandering can take place in a planet either by internal motion²¹ or by the disequilibrium produced by a mountain²² (or crater), provided that solid state creep in the interior can readjust the position of the equatorial bulge which otherwise strongly stabilises the rotation axis. The dates of the magnetised regions on the lunar highlands, which we are discussing, are very early²³: if the rocks are impact breccias they may be the same age as the last great impacts on the Moon (4.1×10^9 yr)²⁵ or, if they arise from craters in the already magnetised highlands crust, they may be older. In any case, they are of such an age that the great impacts which redistributed mass on the lunar surface on a large scale, were a likely cause of the polar wandering required to explain this result. Further studies when a complete magnetic anomaly map is obtained by the lunar polar orbiter may show pole positions distributed between the present equator and the present rotation pole. Such a polar wandering path would be of great interest, and would perhaps be useful for the relative dating of the lunar far side. Such a slow polar wandering through over 90° would suggest that solid state creep in the lunar interior has played an important part in its evolution and supports the

view that convection is responsible for its non hydrostatic shape, as I have maintained^{23,26}. The axis of minimum moment of inertia is aligned by the Earth so that the 'bulge' is earthward: thus if polar wandering has occurred it may have taken place along the present limb so that the ancient axis is now east-west.

I thank P. J. Coleman, S. F. Dermott, L. Hood, R. Lin, L. Molyneux and C. T. Russell for helpful discussions.

S. K. RUNCORN

*Institute of Lunar and Planetary Sciences,
School of Physics, The University,
Newcastle upon Tyne, UK*

Received 22 June; accepted 8 August 1978.

- Hood, L., Russell, C. T. & Coleman, P. J., Jr *Geophys. Res. Lett.* **5**, 305–308 (1978).
- Runcorn, S. K. *et al. Geochim. cosmochim. Acta* **1**, 2369–2387 (1970).
- Mihalov, J. D., Sonett, C. P., Binsack, J. H. & Moutsoulas, M. D. *Science* **171**, 892–895 (1971).
- Coleman, P. J., Jr & Russell, C. T. *Phil. Trans. R. Soc. A* **285**, 489–506 (1977).
- Lin, R., Anderson, K. A., Bush, R., McGuire, R. E. & McCoy, J. E. *Geochim. cosmochim. Acta* **7**, 2691–2703 (1976).
- Dyal, P., Parkin, C. W. & Casson, P. *Geochim. cosmochim. Acta* **3**, 2287–2307 (1972).
- Stephenson, A., Runcorn, S. K. & Collinson, D. W. *Geochim. cosmochim. Acta* **6**, 3049–3062 (1975).
- Taylor, S. R. *Lunar Science: A Post-Apollo View* (Pergamon, Oxford, 1975).
- Runcorn, S. K. *et al. Proc. R. Soc. A* **325**, 157–174 (1971).
- Runcorn, S. K., Libby, L. M. & Libby, W. F. *Nature* **270**, 676–681 (1977).
- Fuller, M. *Rev. Geophys. Space Phys.* **12**, 23 (1974).
- Hide, R. *Moon* **4**, 39 (1972).
- Martelli, G. & Newton, G. *Nature* **269**, 478–480 (1977).
- Gold, T. & Soter, S. *Planet. Space Sci.* **24**, 45 (1976).
- Srnka, L. J. *Geochim. cosmochim. Acta* **8**, 785–792 (1977).
- Strangway, D. W., Rylaarsdam, J. C. & Annan, A. P. *Geochim. cosmochim. Acta* **6**, 2975–2984 (1975).
- Russell, C. T., Coleman, P. J. & Schubert, G. *Science*, **186**, 825–826 (1974).
- Runcorn, S. K. *Phys. Earth Planet. Int.* **10**, 327–335 (1975).
- Creer, K. M., Irving, E. & Runcorn, S. K. *J. Geomag. Geoelect.* **4**, 163–168 (1954).
- Runcorn, S. K. *Ann. de Geophys.* **11**, 73–79 (1955).
- Runcorn, S. K. Giedenkbock, F. A. *Vening Meinesz*, 271–277 (1957).
- Gold, T. *Nature* **175**, 526 (1955).
- Wasserburg, G. J., Papanastassiou, D. A., Tera, F. & Huneke, J. C. *Phil. Trans. R. Soc. A* **285**, 7–22 (1977).
- Turner, G. *Phil. Trans. R. Soc. A* **285**, 97–103 (1977).
- Runcorn, S. K. *Nature* **195**, 1150–1151 (1962).
- Runcorn, S. K. *Proc. R. Soc. A* **296**, 270–284 (1967).

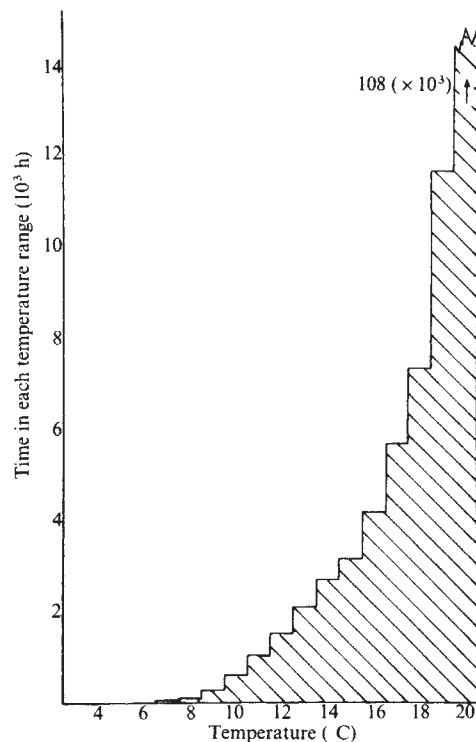


Fig. 1 Histogram of the number of hours spent at different internal temperatures for a 17-yr period (149×10^3 h).

For the main analysis a model system has been constructed in which electrically connected turbines at three typical well-spaced sites (Aberporth, Kinloss and Manston—three of the eight sites considered by Leicester *et al.*²) are used to provide the sole means of heating three groups of houses, one at each site. Each house is equipped with a thermal store having a half-power decay time of 150 h and the heat loss from the house is taken to be proportional to the difference between internal and external temperatures. (Although heat losses also depend to some extent on cooling by wind and rain, these effects are small and have not been included.) For this simple system the rate of supply of heat to the houses is determined by the natural heat loss from the store, although in a practical system some additional control is likely to be included; the houses themselves have a thermal time-constant of 10 h.

Each turbine has a power rating such that the total energy produced over the winter months (October–March) is just sufficient to maintain the houses at that site at a constant internal temperature of 20°C. Because in the short term wind power and heat losses (that is, external temperature) are uncorrelated, there will be some periods when insufficient power is available so that the house temperature will fall and, conversely, others when the supply of heat from the store is excessive and some has to be discarded. To simulate a national grid system the power from the three turbines is divided in such a way as to keep the internal temperatures of the three groups of houses equal and as close to the design value (20°C) as possible.

Table 1 Adopted turbine characteristics^{4,5}

Wind speed		Relative power
Rated wind speed		
0-0.5		0
0.6		0.2
0.75		0.5
0.9		0.8
1.0-2.0		1.0
>2.0		0

Short-term storage and wind power availability

IN a discussion on the economics of alternative energy sources, Ryle¹ suggested that wind power in conjunction with short-term thermal storage can provide a viable and attractive source of energy in the UK. Leicester, Newman and Wright² have recently analysed wind data and empirically derived heating demand over a six-month period. On the basis of these data they suggest that short-term (~150 h) storage is inadequate to provide a reasonably reliable supply from wind generators for heating purposes. This conclusion has been criticised by Diesendorf and Westcott³ and, in view of the limited data sample used it is clear that a more extensive analysis is required. Here we use hourly wind and temperature measurements over a 17-yr period (data obtained from the Meteorological Office, Bracknell) to investigate the performance of a simple system in which wind power is used in conjunction with 150-h thermal storage to provide domestic space heating. The results are promising, even for this basic system, and confirm that, with some modifications, such a system can provide adequate reliability. A significant result of this analysis has been to indicate the importance of selecting an appropriate value for the wind speed at which the wind turbine is to produce its peak output (the 'rated' wind speed). It is shown that the use of too high a rated speed not only greatly exaggerates the fluctuations in available power but could also result in a more expensive design of turbine for a given annual energy output.

The results of this analysis are shown in Fig. 1 as a histogram of the number of hours spent at different internal temperatures. It is found that the temperature is within 3°C of the nominal value (that is $>17^{\circ}\text{C}$) for 86% of the time and falls below 10°C for a total of only 462 h in 17 yr (0.3%). It is interesting to examine the results of our analysis for 1 February to 31 March 1975, a period considered as an 'extreme' example by Leicester *et al.*² The histogram for this period is shown in Fig. 2 (solid line) from which it can be seen that, even over this period, the temperature is above 16°C for over half of the time and never falls below 11°C . (Note that, with no change in the results, the system could also provide domestic hot water at a nominal temperature of, say, 60°C , with corresponding temperature fluctuations.) On the basis of these figures the reliability of the system would not seem to be very much worse than that of the conventional grid when the effects of miners' strikes, 3-day working weeks and so on are included.

The system considered here is extremely simple and possible enhancements which would undoubtedly improve the performance still further are: (1) the provision of a limited amount

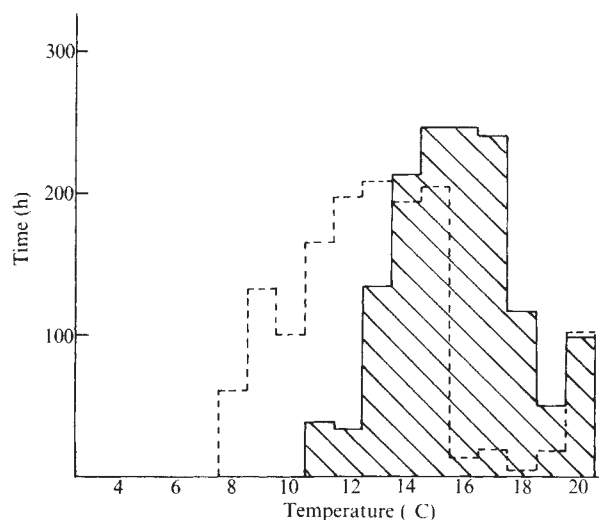


Fig. 2 Histogram corresponding to Fig. 1 for the 'extreme' period, February–March 1975 (Leicester *et al.*²). $V_R = 1.5 \bar{V}$ (solid line); $V_R = 2.3 \bar{V}$ (dotted line).

of non-degrading storage, for example, of the eutectic or chemical type which, in conjunction with the above system, would reduce the wastage of energy during warm periods. (2) The use of storage having the same capacity but better thermal insulation (as might be applicable to a small-scale district storage system) so that some control could be exercised over the rate at which heat was extracted. Avoiding the unnecessary heat loss at night, for example, would be important when considering the economics of the system.

We have used the measured performance figures of actual machines^{4,5} and, as described below, adopted a 'rated' wind speed of 1.5 times the mean wind speed, $\bar{V}$. By comparison, the generator characteristics used by Leicester *et al.*² are based on a higher rated wind speed and, also, give an output which seems to fall off unreasonably rapidly at the lower wind speeds. To show the effect of a higher rated speed we have recomputed the histogram of Fig. 2 for a rated speed of $2.3 \bar{V}$ (dotted line) from which it is apparent that the performance of the system is significantly degraded. In contrast to the results of Leicester *et al.*², who assumed an unrestricted rate of energy extraction from the store, in neither case is the store completely emptied over this period. The discrepancy between our results and those of Leicester *et al.*² can therefore be explained largely in terms of (1) the different wind-turbine characteristics which we have used

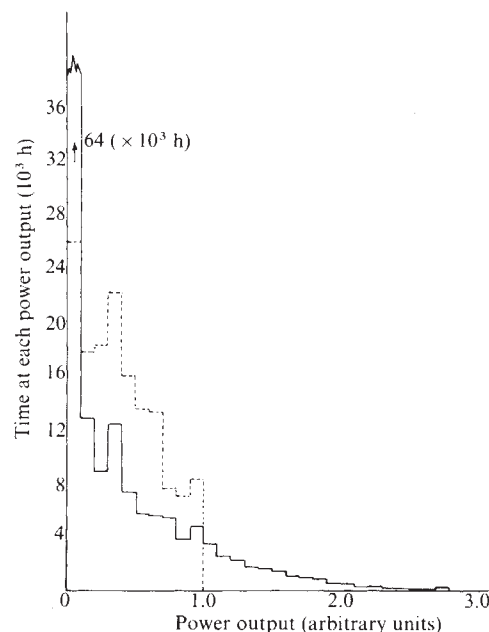


Fig. 3 Temporal distribution of combined power output from the three sites for two values of 'rated' wind speed; $V_R = 2.3 \bar{V}$ (solid line); $V_R = 1.5 \bar{V}$ (dotted line).

(Table I) and (2) our use of a thermal store with an exponential decay. (However, it is not clear why the averaged heating demand as derived by Leicester *et al.*² should have such a low standard deviation (4.9%), a value which is remarkably small even allowing for a significant, and constant, demand for domestic hot water; in contrast, the weekly records of the energy used for heating our own laboratory between October 1977 and March 1978 show a s.d. of 30%.)

Finally, we discuss briefly the effect of variations in the rated wind speed of the turbine. ETSU⁶ have shown that for UK sites the maximum annual energy output is provided by a turbine having a rated wind speed approximately 2.3 times the mean wind speed, assuming that the cut-in speed, that is the speed below which no output is obtained, is about half the rated wind speed. However, such a turbine gives rise to prolonged gaps in the energy supply. If a lower rated wind speed is adopted these gaps are considerably reduced; conversely, if the rated speed is higher than $2.3 \bar{V}$, not only are the gaps seriously increased but the total energy output also falls. The effect of changes in the rated speed can be seen clearly in Fig. 3 which shows the number of hours for which a given combined output is obtained from the three sites considered previously. The histograms have been derived from two systems having the same total energy output

Table 2 Parameters of two wind turbines having the same annual output but different rated speeds

Parameters	$V_R = 2.3 \bar{V}$	$V_R = 1.5 \bar{V}$
Power rating	1 MW	300 kW
Rated wind speed	27 knots	18 knots
Diameter	46 m	52 m
Relative axial force	1.0	0.54
Relative torque	1.0	0.62
Estimated costs (£10 ³)		
Tower, foundations	22	17
Orientation	4	4
Rotor	30	34
Mechanical transmission	26	16
Alternator	12	5
Other items	14	14
Total cost	£108(000)	£90(000)

but with different rated speeds of $2.3 \bar{V}$ (full line) and $1.5 \bar{V}$ (dotted line). It is apparent that the former system has much longer periods of low or zero output and a correspondingly shorter time at which it is near its rated output.

Because, for a given swept area, reducing the rated speed from $2.3 \bar{V}$ to $1.5 \bar{V}$ decreases the annual energy output by ~20%, a proportionate increase has to be made to the size of the turbine blades to obtain the same mean output. The latter machine is, nevertheless, likely to be cheaper, because of the reduced ratings of the alternator and gearbox and the lower loading of the tower, as well as providing a more steady output. This is shown in Table 2 in which two designs having the same annual output are compared. The relative costs have been estimated using figures in the ETSU report⁶.

Our analysis has been concerned only with the time variations of available wind power in relation to fluctuations in the demand for space heating. Because of the normalisation carried out at the start of the analysis no knowledge of the absolute power demand is required and the results are independent of the type or number of houses to be supplied by the wind turbine or the precise temperature at which they are maintained. However, an approximate estimate of absolute energy requirements can readily be made. Assuming an average winter heating demand of 2 kW, a typical figure for a medium-sized, well-insulated house, a single turbine of the type considered in Table 2 would be sufficient to maintain about 90 houses at the temperature levels given in Fig. 1.

We thank the Meteorological Office for the supply of wind data. Financial assistance from the SRC to M.B.A. and from the Royal Commission for the Exhibition of 1851 to K.N. is acknowledged.

M. B. ANDERSON
K. NEWTON
MARTIN RYLE
P. F. SCOTT

Cavendish Laboratory,
Madingley Road,
Cambridge, UK

Received 26 June; accepted 24 August 1978.

1. Ryle, M. *Nature* **267**, 111-117 (1977).
2. Leicester, R. J., Newman, V. G. & Wright, J. K. *Nature* **272**, 518-521 (1978).
3. Diesendorf, M. & Westcott, M. *Nature* (in the press).
4. Thomas, T. L. *Vertical Axis Wind Turbine Technology Workshop*, New Mexico, May (1976).
5. Hütter, U. *Operating Experience Obtained with a 100kW Wind Plant* (NASA, TT-F-15, 68, 1973).
6. Energy Technology Support Unit, Dept of Energy, *Energy Paper* **21** (HMSO, London, 1977).

New materials in the Si-C-Al-O-N and related systems

THE concurrent discovery at the Universities of Utah and Newcastle-upon-Tyne of a continuous series of solid solutions, with the 2H wurtzite-type structure, between silicon carbide (SiC) and aluminium nitride (AlN) is reported here. This, together with observations of the additional solubility of aluminium oxycarbide (Al_2OC), suggests a large extension to the already wide field of sialons. It offers the prospect of many more materials in the Si-C-Al-O-N and related systems that are of scientific and technological interest not only as engineering ceramics but also for their physical, chemical, and electrical properties.

Silicon nitride and silicon carbide are leading contenders for gas turbine and other high-temperature engineering applications. Extensive solid solution based on $\beta\text{-Si}_3\text{N}_4$ was first shown by Oyama¹ and by K.H.J. and Wilson² but only a few systems, such as $\text{Si}_3\text{N}_4\text{-Al}_2\text{O}_3\text{N}$, $\text{Si}_3\text{N}_4\text{-MgAl}_2\text{O}_4$ and $\text{Si}_3\text{N}_4\text{-}$

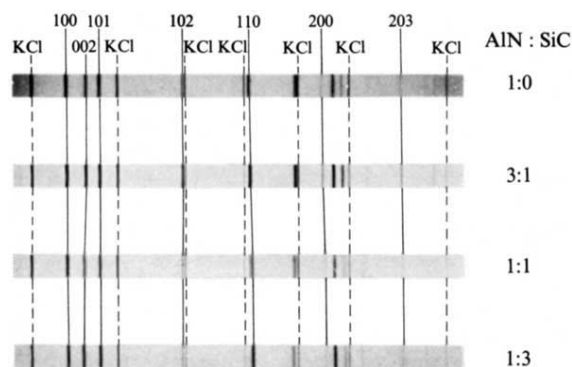


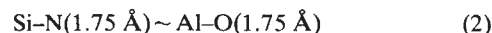
Fig. 1 X-ray photographs of 2H wurtzite-type solid solutions of SiC-AlN.

LiAl_5O_8 , have been explored in detail. Investigation progressed in different directions³ when it was appreciated that the reversible replacement

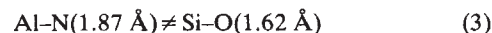


is of general applicability and that the sialons are essentially aluminosilicates in which oxygen is partly replaced by nitrogen. A wide field of new materials exists, vitreous as well as crystalline, built up of $(\text{Si,Al})(\text{O,N})_4$ tetrahedra in the same way that the almost infinite variety of mineral silicates is based upon the $(\text{Si,Al})\text{O}_4$ tetrahedron.

The formation of extensive solid solutions with silicon nitride and also with silicon carbide might seem surprising in that their covalent bonding, in addition to its high strength, is specific and selective. In the Si-Al-O-N system³ the extended homogeneity range of phases along lines of constant metal:nonmetal atom ratio M:X and the very restricted solubility in directions parallel to the AlN-SiO₂ join is due, at least in part, to the similarity of the two bond lengths



and the difference between the other two bonds



The ready replacement of metal and nonmetal atoms in covalently bonded solids is not restricted, however, to that of Si-N by Al-O. Thus, a variety of polytypes M_mX_{m+1} , $M_{m+1}X_m$ and MX with extended regions of homogeneity have been characterised recently⁴ in the Mg-Si-Al-O-N, Be-Si-Al-O-N and Al-C-O-N systems. Their structures are related to those of the silicon carbide polytypes and are based on the atomic arrangement of wurtzite.

Recent work at Newcastle⁵ has shown the existence of α' -sialons, that is, phases in the M-Si-Al-O-N system where M includes Li, Mg, Ca and Y, that are isostructural with $\alpha\text{-Si}_3\text{N}_4$. The unit-cell contents are represented by $M_x(\text{Si,Al})_{12}(\text{O,N})_{16}$ where $x \geq 2$ and where the metal atoms M occupy partially or completely the only two interstitial sites in the structure. As

Table 1 Unit-cell dimensions of the end members of wurtzite-type solid solutions

	Dimensions (Å)	
	a	c
$\alpha\text{-SiC}(2\text{H})$	3.076	5.048
AlN	3.114	4.986
Al_2OC	3.19	5.09

shown by equations (4) and (5), the formation of α' -sialons from silicon nitride



is essentially by replacement of Si-N by Al-N and not by Al-O as in β' formation.

Although the preparation⁶ of wurtzite-type Al_3CON shows the possibility of producing carbon-containing sialons, the first attempts to do so at Newcastle were unsuccessful. Because diffusion coefficients in covalent solids are extremely small, solid solution is unlikely to be obtained by the heating and annealing of the powdered solid components; nor is it possible by melting. The other possibilities, by liquid-phase and vapour-phase reactions, have both now been used in preparing SiC-AlN solid solutions.

The preparation of silicon carbide and silicon nitride from silica by reduction with carbon or carbon plus nitrogen is a vapour-phase process⁷⁻⁹ involving the formation of SiO and CO. Similar reactions occur in the production of aluminium nitride from aluminium oxide (I.B.C., unpublished) and of sialons from clay¹⁰ and $\text{SiO}_2 : \text{Al}_2\text{O}_3$ gels (A. Hendry and K.H.J., unpublished). All such processes occur rapidly at relatively low temperatures, 1,400–1,600 °C. By using a very finely divided, high surface area, amorphous silica ('Cabosil') and a carbon source such as starch or sugar, an intimate and porous mixture of carbon and silica is produced by coking to remove all volatiles. Other oxides may then be homogeneously dispersed within the mixture by precipitation from their soluble salts. Using these techniques, with aluminium precipitated as hydroxide, reactions have been carried out at Utah in both argon and nitrogen atmospheres at 1,600 °C. These show 2H wurtzite-type solid solutions between SiC and AlN and between SiC and Al_2OC extending in each case from 2m/o to 100m/o of the aluminium-containing component. In each case continuous shifts of reflections were observed in a single-phase wurtzite-type

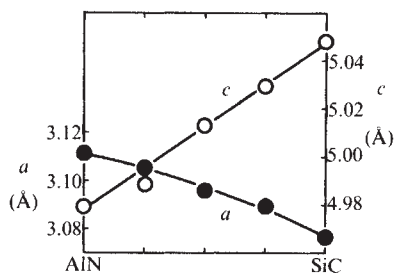


Fig. 2 Unit-cell dimensions of 2H solid solutions of SiC-AlN.

diffraction pattern corresponding with the unit-cell dimensions for the end members given in Table 1.

At Newcastle it was considered that α' -sialons were obtained by reaction of Si_3N_4 with AlN plus relatively small concentrations of CaO or Ca_3N_2 because of the formation of a liquid phase in which the reactants have some solubility. When α' - $\text{CaSi}_9\text{Al}_3\text{ON}_{15}$ was reduced in a carbon crucible at 1,800 °C for 30 min in flowing nitrogen the product was single phase with a wurtzite-type structure of unit-cell dimensions corresponding to 3SiC . AlN. No calcium was detected in the product and in the same experimental conditions calcium oxide was completely reduced and volatilised as calcium vapour and carbon monoxide. In subsequent experiments at the same temperature, mixtures of Si_3N_4 , AlN and CaO were reacted with carbon instead of using the preformed α' -sialon. It is thought that the addition of CaO provides a transient oxynitride liquid in which the reactants are soluble.

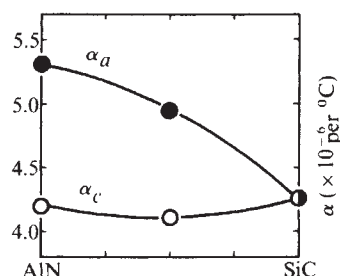


Fig. 3 Linear coefficients of expansion (20–1,000 °C) of solid solutions of SiC-AlN.

SiC-AlN solid solutions were also obtained at Newcastle by heating intimate mixtures in molecular nitrogen at 1,400–1,500 °C of SiO_2 , Al_2O_3 and C obtained by a sol-gel process. The products always contained some β -SiC, unlike those prepared by essentially the same process at Utah. The formation of β -SiC depends on impurities, and also on nitrogen pressure, but it is clear from the present results that β , once formed, does not readily react with AlN to give a solid solution.

Figure 1 shows X-ray photographs of 2H solid solutions of 3AlN : SiC, AlN : SiC and AlN : 3SiC compared with that of pure AlN. The shifts of X-ray reflections show that c increases and a decreases as AlN is replaced by SiC. This variation in cell dimensions is shown by Fig. 2. The mean coefficients of expansion α_a and α_c at 20–1,000 °C for the composition AlN : SiC are compared with values for the end members in Fig. 3. The value of about 5×10^{-6} per °C is probably too high for the use of the material in gas turbine applications but the electrical and semi-conducting properties are expected to be of interest.

As extensive solubility between the end members of the SiC-AlN solid solution occurs with Al_2OC , five component Si-C-Al-O-N phases must occur although their ranges of homogeneity have not yet been established. Moreover, there are many nitrides, oxynitrides and carbo-oxynitrides with wurtzite-type structures, for example, LiSi_2N_3 , MgSiN_2 , LiSiON , MgSiAlN_3 and Al_3CON , and so in appropriate conditions these should also form solid solutions with SiC-AlN.

The new materials based on Si-C-Al-O-N in combination with a variety of network-modifying metals will not all be useful ceramics but their other properties are worth investigation. Moreover, they provide a new and hitherto unexplored area of inorganic chemistry.

I. B. CUTLER
P. D. MILLER
W. RAFANIELLO

Department of Materials Science and Engineering,
The University of Utah,
Salt Lake City, Utah 84112

H. K. PARK
D. P. THOMPSON
K. H. JACK

Wolfson Research Group for High-Strength Materials,
Crystallography Laboratory,
The University,
Newcastle-upon-Tyne, UK

Received 4 August; accepted 16 August 1978.

- Oyama, Y. *Jap. J. appl. Phys.* **11**, 760 (1972).
- Jack, K. H. & Wilson, W. I. *Nature phys. Sci.* **238**, 28 (1972).
- Jack, K. H. *J. mater. Sci.* **11**, 1135 (1976).
- Roebuck, P. H. A. & Thompson, D. P. *High Temperature Chemistry of Inorganic and Ceramic Materials* (eds Glasser, F. P. & Potter, P. E.) 222–228 (The Chemical Society, London, 1977).
- Hampshire, S., Park, H. K., Thompson, D. P. & Jack, K. H. *Nature* **274**, 880 (1978).
- Roebuck, P. H. A. thesis, Univ. Newcastle-upon-Tyne (1978).
- Lee, J. G. & Cutler, I. B. *Bull. Am. ceram. Soc.* **54**, 195 (1975).
- Lee, J. G. & Cutler, I. B. *Proc. NATO Adv. Study Inst., Nitrogen Ceram.* 175–181 (Noordhoff, Leyden, 1977).
- Hendry, A. & Jack, K. H. *Special Ceramics* (ed. Popper, P.) **6**, 199–208 (The Brit. Ceram. Res. Assoc., Stoke-on-Trent, 1975).
- Lee, J. G., Casarini, R. & Cutler, I. B. *Paper at the Am. ceram. Soc. Mtg* (October, 1975).

New isotopic evidence for the origin of Red Sea brines

THE hot brines in the Atlantis II Deep in the Red Sea (Fig. 1) have concentrations of both oxygen- and hydrogen-isotopes which are different from those of the overlying Red Sea deep water but are identical to the surface waters, only at the southern parts of the Red Sea^{1,2}. It was inferred, therefore, that these brines originated in the southern Red Sea and have migrated northwards to their present location. This hypothesis has often been questioned, using mainly geological arguments, the most striking of which is the improbability of a plumbing system which had to work over a distance of more than 800 km (refs 3, 4). However, the geochemical problem of the depletion in heavy isotopes in the hot brines as compared to the directly overlying seawater has still not been solved. We can add new data to this discussion from stable isotope analyses carried out on new brines which were found throughout the Red Sea (Fig. 1). Many of these occurrences are between the Atlantis II Deep and the southern Red Sea⁵.

Our analytical procedures were the same as those for the early isotope analyses on the Atlantis II Deep brines¹ and are described in detail elsewhere^{6,13}. The accuracies were $\pm 0.07\%$ and $\pm 0.5\%$ for oxygen and hydrogen, respectively.

The results are given in Table 1 and Fig. 2. Many of the brines have chlorinities which are appreciably lower than the nearly saturated brine in the Atlantis II Deep. We assume that this is due to dilution by seawater after intrusion of the saturated brines. This would have caused a change in the isotopic compositions if seawater and brines differed in their initial isotope content. Therefore, we calculated the assumed compositions of the brines before their dilution. These 'corrected' values are shown in Table 1 and Fig. 2.

The brines exhibit a considerable variation in their isotopic content. However, those brines near the Atlantis II Deep are isotopically identical. These include the Atlantis II Deep itself, the Discovery Deep, the Chain Deeps and the Albatross Deep (Fig. 1). This homogeneity supports previous assumptions^{5,7} that the brines in the Atlantis II Deep area are from one source and

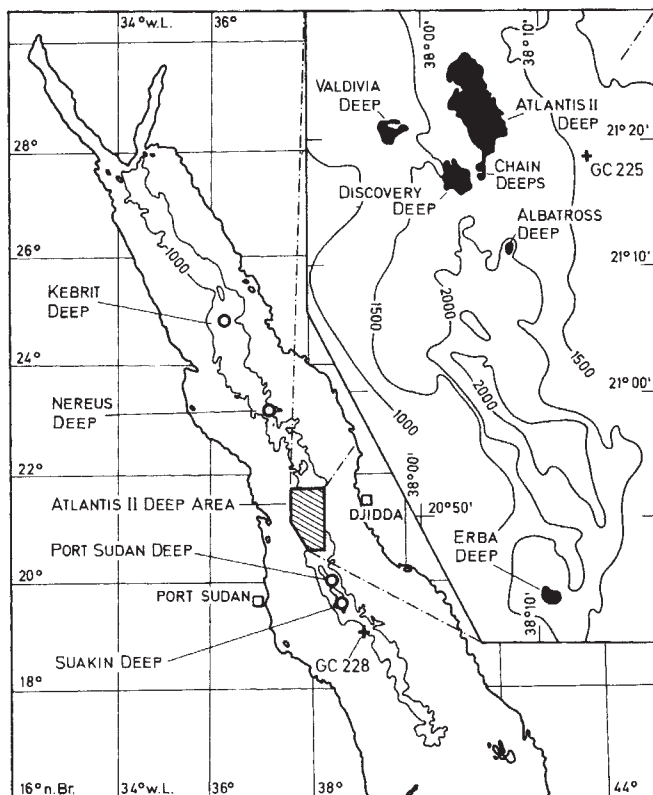


Fig. 1 Locations of Atlantis II Deep brine pools and of new brine occurrences in the Red Sea^{5,19}.

Table 1 $^{18}\text{O}/^{16}\text{O}$, D/H-isotope and chloride data for waters of various brines in the Red Sea

Location	Cl (gperkg)	$\delta^{18}\text{O}^*$ (‰)	δD^* (‰)	$\delta^{18}\text{O}^\dagger$ (‰)	$\delta\text{D}^\dagger$ (‰)
Atlantis II Deep	+156.0	+1.21	+7.4	+1.19	+7.3
Chain A-Deep	+154.8	+1.29	+8.2	+1.26	+8.0
Chain B-Deep	+155.5	+1.19	+7.5	+1.16	+7.3
Chain C-Deep	+154.2	+1.22	+7.5	+1.19	+7.3
Discovery Deep	+155.1	+1.16	+7.5	+1.13	+7.3
Albatross Deep	+143.5	+1.20	+7.7	+1.10	+7.2
Valdivia Deep	+144.6	+1.84	+11.5	+1.83	+11.5
Erba-Deep	+83.8	+1.98	+13.0	+2.09	+14.9
Port Sudan-Deep	+124.8	+2.39	+15.0	+2.57	+16.2
Nereus-Deep	+129.5	+2.64	+17.2	+2.86	+18.9
Kebrit-Deep	+153.2	+1.05	+10.4	+1.00	+10.3
Suakin-Deep	+85.8	+2.56	+21.6	+3.35	+33.6

δ -values are reported versus SMOW and are defined as

$$\delta = \frac{R_{\text{sample}} - R_{\text{standard}}}{R_{\text{standard}}} \times 1,000[\text{‰}]$$

where R is the D/H and the $^{18}\text{O}/^{16}\text{O}$ ratio respectively.

* Measured values. The original brines are supposed to be diluted by Red Sea Water so the corrected data[†] for undiluted brines are calculated according to the following formula:

$$\delta_{\text{corr}}^i = (\delta_{\text{Brine}} - x_i \delta_{\text{SW}}) / (1 - x_i)$$

δ_{Brine}^i are the measured isotope data*, δ_{SW} the data of recent Red Sea deep water ($\delta^{18}\text{O}_{\text{SW}} = 1.89\text{‰}$, $\delta\text{D}_{\text{SW}} = 11.5\text{‰}$ (ref. 6)). The sea water contributions x_i to the actual brines are calculated from the respective contents $\text{Cl}_{\text{Brine}}^i$: $x_i = (\text{Cl}_{\text{Brine}}^i - 160.8) / (22.5 - 160.8)$. The saturation Cl^- -concentration of Atlantis II brine is 160.8 g per kg at 25 °C (ref. 2); 22.5 g per kg (ref. 6) was determined for recent Red Sea deep water.

probably have subsurface connection. The presence of nearly identical brines in the Albatross Deep is now understood better and could not be explained by the overspill theory⁸.

The Valdivia Deep is an elevated position located away from the central rift 10 km west of the Atlantis II Deep system (Fig. 1). Its brines have significantly higher D/H and $^{18}\text{O}/^{16}\text{O}$ ratios than the Atlantis II Deep brines. They are the only example of brines which are isotopically identical to present-day Red Sea water, which suggests that they originated from present-day Red Sea deep water, or from Red Sea palaeowaters of present-day composition. For this brine pool, a dissolution of exposed evaporites⁸ by present day Red Sea deep waters is likely.

The brines from the Erba, Port Sudan and Nereus Deeps are enriched in both D and ^{18}O and exceed the range of present-day Red Sea deep waters. Only a few processes could have caused a simultaneous enrichment of D and ^{18}O as shown in Fig. 2 for these three brines. An enrichment by evaporation is highly unlikely as these brines have never been exposed to air. Mixing of recent Red Sea water with only one subsurface brine having a specific isotopic composition is excluded as the adjustment of these values to brines unmixed with seawater (corrected values in Table 1) should then result in identical isotopic compositions, and the corrected values differ significantly. The δD and $\delta^{18}\text{O}$ values of the three brines under discussion fall nearly on a line in the $\delta\text{D}/\delta^{18}\text{O}$ diagram; this line is located between present-day Red Sea deep water and Red Sea deep water of the last glacial maximum 18,000 yr BP (III in Fig. 2). This value was estimated using ^{18}O -variations in foraminifera in Red Sea sediments which are given in Fig. 3. The total ^{18}O -enrichment, $\Delta^{18}\text{O}$, during the last glacial maximum was $\sim +2\text{‰}$ in ^{18}O as compared with present-day values. The $\Delta^{18}\text{O}$ values in Fig. 3 are the approximate ^{18}O -oscillations of the Red Sea waters during the past 20,000 yr as surface water temperatures did not change significantly during this period¹⁰. The Red Sea deep water-values given in Fig. 2 have been estimated with the assumption that the $\delta^{18}\text{O}/\delta\text{D}$ relationship of Red Sea deep water relative to open ocean water (SMOW) (that is $\delta\text{D} = 6\delta^{18}\text{O}$ (ref. 1)) has prevailed.

We conclude that the brines from below the Erba, Port Sudan and Nereus Deep are Red Sea waters, which infiltrated to sub-bottom strata during times when the isotopic composition of the Red Sea was significantly enriched in heavy isotopes as compared with today; that is, these brines are essentially derived from Red Sea palaeowaters.

The same concept could also explain the origin of the Atlantis II Deep brines. During past pluvial periods the Red Sea water was depleted in ^{18}O by up to 1‰ compared to open ocean water¹¹. An ^{18}O -depletion is also indicated in Fig. 3 by the negative $\Delta^{18}\text{O}$ values at inferred ages that are about 5,000 yr BP. We therefore propose that the isotopic composition of the Atlantis II Deep should not be related to present-day Red Sea surface waters but more likely to isotopically depleted Red Sea deep waters from previous pluvial periods. The depletion in ^{18}O of $\sim 0.5\%$ during the last climatic optimum ($\sim 5,000$ – $8,000$ yr BP) as indicated in Fig. 3 is sufficient to explain the isotopic composition of the Atlantis II Deep brines which are also depleted by 0.6‰ as compared to present day Red Sea deep waters.

The concept of a palaeowater origin of the brines in the Red Sea agrees with the widespread occurrence of brines and eliminates the need to hypothesise their derivation from waters to the far south. We assume that these palaeowaters were trapped in an earlier geological period and have been stored since then in subterranean strata. Alternatively, dynamic systems have to be assumed where the subsurface flow times of these systems are of the order of 7,000–15,000 yr.

The brines in the Kebrit and Suakin Deeps do not have isotopic compositions within the range of possible Red Sea palaeowater compositions. Some specific waters of the meteoric water cycle have similar compositions. Such waters were found in lakes in the Danakil Desert¹² or in hot springs in the area around Djibouti¹³. The intrusion of such waters at the bottom of the Red Sea is unrealistic. Therefore, we assume the Kebrit and Suakin brines to be Red Sea palaeowaters which have changed their isotopic compositions due to secondary effects. High temperature exchange reactions between minerals and waters

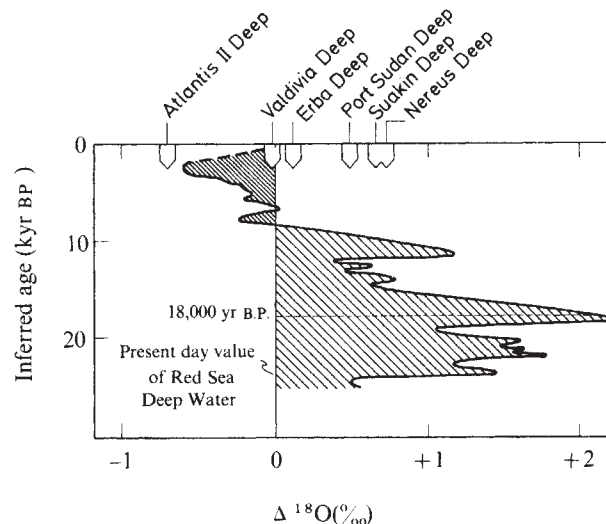


Fig. 3 Variation of the ^{18}O concentration in the Red Sea during the past 20,000 yr. The $\Delta^{18}\text{O}$ values are defined as $\delta^{18}\text{O}_{\text{past}} - \delta^{18}\text{O}_{\text{recent}}$ and show the relative change of ^{18}O in foraminifera in the Red Sea as compared to present day foraminifera values. As the temperature variations in the Red Sea waters were small, the $\Delta^{18}\text{O}$ values represent approximately the variations in ^{18}O in sea water. The time scale has been inferred by assuming the time of maximum ^{18}O enrichment to be 18,000 Yr BP. The ^{18}O concentrations of the brines are given for comparison and are also plotted as differences from present day Red Sea deep water. This comparison shows that the variations in ^{18}O in seawater cover those in the brines.

are improbable. Even the Atlantis II Deep brines, which exhibit the highest temperatures, do not show any ^{18}O isotope shift². We suggest, however, that low temperature mineral reactions, similar to those which alter pore fluid compositions^{14,15} should be considered. Interstitial waters in the Red Sea sediments have been analysed in bore holes near the brine deeps (GC 225 and GC 228 in Fig. 1). Although the ^{18}O concentrations in these fluids could be determined only with great uncertainty (J. R. Lawrence, personal communication), it is important to our results that interstitial waters are similar in their isotopic compositions to those brines which do not fit to proposed unaltered palaeowaters. These waters also tend to be different as compared to Red Sea deep waters and Red Sea palaeowaters and are similar to the brines at the Suakin and Kebrit Deeps. The brines of the Kebrit and Suakin Deep may, therefore, be interpreted as accumulations of expelled interstitial waters.

This interpretation is somewhat speculative, especially taking into account the isotopic composition of the Suakin brines before dilution. However, alternative source fluids like downwards migrated meteoric waters from the Saudi Arabian coast³ are even more speculative and formation waters¹⁸ are ruled out by their isotopic composition.

We thank E. T. Degens for helpful discussions and C. E. Redding for reading the manuscript. This research was funded by Bundesministerium für Bildung und Wissenschaft.

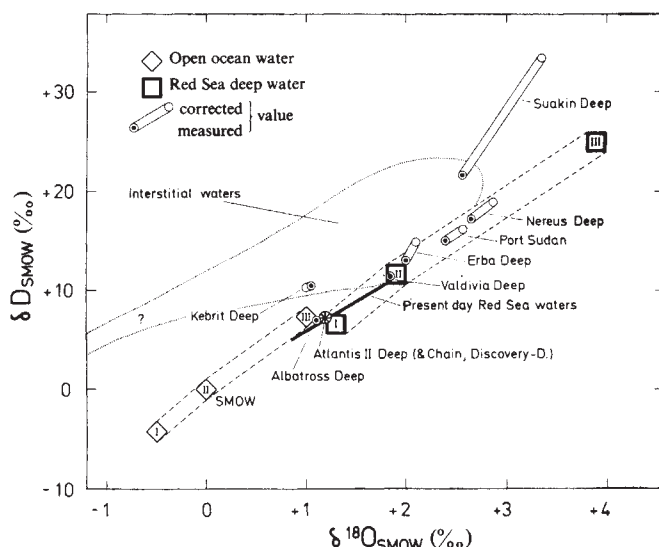
M. SCHOELL
E. FABER

Bundesanstalt für Geowissenschaften und Rohstoffe,
Postfach 51 01 53,
3000 Hannover 51, FRG

Received 15 May; accepted 17 July 1978.

1. Craig, H. *Science* **154**, 1544 (1966).
2. Craig, H. in *Hot Brines and Recent Heavy Metal Deposits in the Red Sea* (eds Degens, E. T. & Ross, D. A.) (Springer, Berlin, 1969).
3. White, D. E. *Econ. Geol.* **63**, 301 (1968).
4. Manheim, F. T. *Init. Rep. DSDP* **23**, 975 (1974).
5. Bäcker, H. & Schoell, M. *Nature Phys. Sci.* **240**, 153 (1972).
6. Faber, E. *Bericht Bundesanstalt für Geowissenschaften* (unpublished).
7. Schoell, M., Bäcker, H. & Baumann, A. in *Problems of Ore Deposition*, 4th IAGOD Symp. **1**, 303 (1974).

Fig. 2 Isotopic composition of Red Sea brines and some inferred source fluids. Each brine which has a salinity higher than the Atlantis II Deep brine has a corrected value (○) which is connected with the measured value (●). The corrected values give the original isotopic composition of the brines, if it is assumed that the lower salinities were caused by dilution with present-day seawater. The isotopic variation of world ocean water and Red Sea deep water is shown for the following stages: I interglacial (last climatic optimum 8,000–5,000 yr BP), II present day and III glacial maximum 18,000 yr BP. The isotopic composition of interstitial waters is not well defined (J. R. Lawrence, personal communication); the indicated field encloses the reported analytical uncertainties¹⁶.



8. Ross, D. A. *et al. Science* **179**, 377 (1973).
9. Schoell, M. & Risch, H. *Geol. Jb. D* **17**, 15 (1976).
10. McIntyre *et al. Science* **191**, 1131 (1976).
11. Deuser, W. G., Ross, E. H. & Waterman, L. S. *Science* **191**, 1168 (1976).
12. Gonfiantini, R., Borsi, S., Ferrara, G. & Pianchi, C. *Earth planet. Sci. Lett.* **18**, 13 (1973).
13. Schoell, M. & Faber, E. *Geol. Jb. D* **17**, 197 (1976).
14. Perry, E. A., Gieskes, J. M. & Lawrence, J. R. *Geochim. cosmochim. Acta.* **40**, 413 (1976).
15. Lawrence, J. R., Gieskes, J. M. & Broecker, W. S. *Earth planet. Sci. Lett.* **27**, 1 (1975).
16. Lawrence, J. R. *Init. Rep. DSDP* **23**, 939 (1974).
17. Friedman, I. & Hardcastle, K. *Init. Rep. DSDP* **23**, 969 (1974).
18. Issar, A., Rosenthal, E., Eckstein, Y. & Bogoch, R. *Int. Ass. Scient. Hydrol.* **16**, 39 (1971).
19. Bäcker, H., Lange, K. & Richter, H. *Geol. Jb. Reihe D, Heft* **13**, 79–123.

Texture discrimination and Fourier analysis in human vision

NEUROPHYSIOLOGICAL experiments have shown that many cells in the visual cortex of the cat and monkey respond well only if the retinal image contains a line stimulus of appropriate orientation and width (spatial frequency)^{1,2}. Different cells are sensitive to different orientation/spatial frequency combinations and judging from numerous corroborative psychophysical experiments, it seems highly probable that similar populations of cells exist in man also³. However, the function of these cells is not known. One suggestion is that they decompose the retinal image into its two-dimensional Fourier components and that processes such as object recognition, region finding and image segmentation then operate on this Fourier description^{4–7}. We report here experimental results on human texture discrimination which argue against this possibility by showing that textures which differ markedly in their Fourier spectrum are not always readily discriminable.

Figure 1 illustrates the texture discrimination task we used. Each stimulus had one 'target' quadrant whose texture differed from that used for the other three quadrants. The stimuli were presented in a tachistoscope and latencies were measured for discrimination of the target. A buzzer warned the subject when a stimulus was about to appear and he indicated his choice for target quadrant by pressing one of four response keys, arranged in a panel to match the layout of the stimulus quadrants. His response terminated the stimulus presentation and also stopped a timer which recorded response latency. Each stimulus was presented four times during an experimental session lasting about 30 min. The location of the target quadrant was varied randomly from trial to trial except that each quadrant was used once for the target for each stimulus. Stimuli were presented in a different quasi-random order for each subject which ensured that he did not know which quadrant was to be the target on any given trial.

Textures consisted of either one sinusoidal grating (45°, Fig. 1a and d), two sinusoidal gratings added together (45° + 105°: Fig. 1b and e), or three sinusoidal gratings added together (45° + 105° + 165°: Fig. 1c and f). In half the stimuli the target quadrant differed from the other quadrants in the spatial frequency of its texture (2.2 and 4.4 cycles per degree respectively: Fig. 1a–c). In the other half the target differed in the orientation of its components (30° rotation: Fig. 1d–f). All stimuli were generated by computer and produced as hardcopy on a full-tone facsimile picture recorder (Muirhead type K-550-B/1).

The reader is invited to judge for himself the difficulty of finding the target quadrant in each of the stimuli shown in Fig. 1. Subjects made the various spatial frequency discriminations very quickly but took much longer over the orientation ones. Even the quickest orientation case (Fig. 1d) produced a mean latency twice as long as any of the spatial frequency ones and in the three-component orientation condition (Fig. 1f) the latency is extremely long indeed.

We have also studied this kind of texture discrimination task using 200-ms tachistoscopic stimulus presentations which prevent benefit from scanning eye movements. Texture discriminations which can be made in such circumstances are

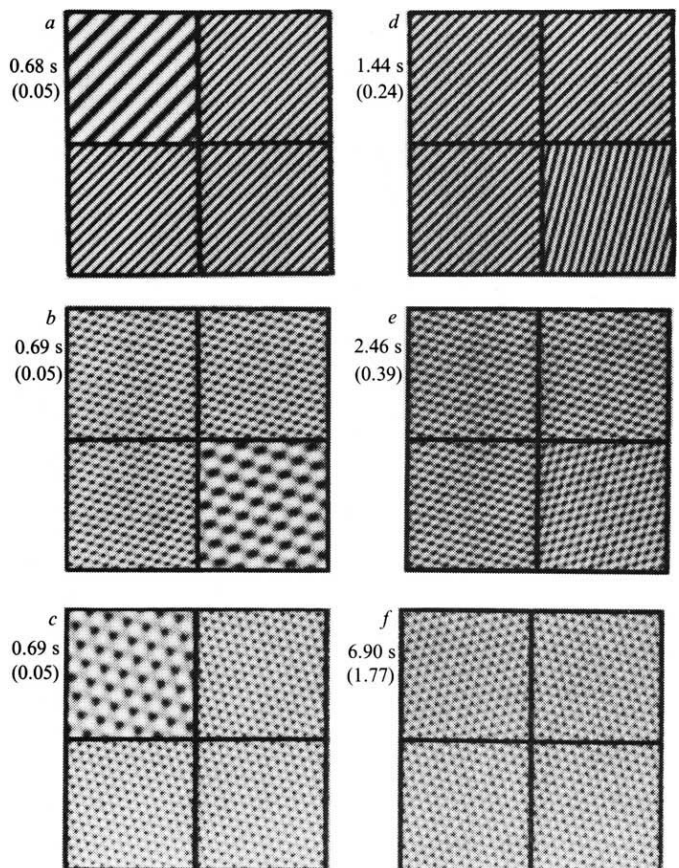
sometimes called 'immediate' to distinguish them from discriminations requiring much more lengthy and careful scrutiny⁸. The spatial frequency discriminations remained easy using 200 ms presentations whereas the orientated type could not be performed reliably, even in the one-component condition⁹. We have also found that these conclusions apply to textures composed of filtered random noise (which appear like woven fabrics).

These results are difficult to reconcile with any theory in which it is assumed that the processes of texture discrimination have immediate access to information corresponding to a local piecewise Fourier analysis. In Fourier terms, the difference between target and non-target textures in all the stimuli of Fig. 1 is clearcut, consisting as they do of widely separated pure sinusoids. Given a Fourier model, therefore, easy discrimination would be expected in all cases but this was not the obtained result.

The results are also difficult to reconcile with Julesz's theory of texture vision¹⁰. Julesz has proposed that textures which differ only in their 3rd or higher order statistics cannot be discriminated without scrutiny. This conjecture does not entail, of course, that all textures differing in their lower order statistics be discriminable. Even so, the second order differences between target and non-target quadrants in Fig. 1 are marked and this makes the relative difficulty with orientated discriminations surprising, given Julesz's basic thesis.

An alternative theory of texture vision has been formulated by Marr¹¹ in which the processes of texture vision have access to a rich but primitive symbolic description of the visual scene called the 'primal sketch'. According to Marr's theory, the primal sketch is computed from local measurements taken at every position in the visual image by the orientated spatial frequency channels. The outputs of these channels are then used to produce a description of the kind of edge or blob present at

Fig. 1 Sample stimuli used in the experiment. Times alongside each figure are group mean latencies ($n = 9$) for detection of the target quadrant (s.e.m. in brackets).



each location, with channel outputs not available to subsequent processes. Consider, for example, Fig. 1c. In terms of Marr's theory, the measurements taken by the various orientated channels are used to form 'assertions', entered into the primal sketch, about the presence of blobs of a particular size and contrast. The clear cut differences in orientated spatial frequency channel activity which would be generated are of no relevance to subsequent processes. Instead, in order to detect the odd quadrant higher-order grouping processes are required to aggregate the blobs into rows. It is only after this grouping process (which would be maximally difficult in Fig. 1f—hence the longest latency for this condition) that the orientation differences in the textures can be detected. On the other hand, in the case of a discrimination based on spatial frequency differences, no such processing is required as the size of the blobs is written directly into the primal sketch and is thus available directly to the processes of immediate texture discrimination. Clearly, our results are compatible with Marr's ideas and challenge any theory which suggests that the outputs of orientated spatial frequency channels are used directly for immediate texture vision.

J. E. W. MAYHEW
J. P. FRISBY

Department of Psychology,
University of Sheffield,
Sheffield, UK

Received 9 December 1977; accepted 25 August 1978.

- Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **160**, 106–154 (1962); **165**, 559–568 (1963); **195**, 215–243 (1968); *J. comp. Neurol.* **58**, 267–294 (1974).
- Hubel, D. H., Wiesel, T. N. & Stryker, M. P. *Nature* **269**, 328–330 (1977).
- Mostafavi, H. & Sakrison, D. J. *Vision Res.* **16**, 957–968 (1976).
- Blakemore, C. & Campbell, F. W. *J. Physiol., Lond.* **203**, 237–260 (1969).
- Pollen, D. A., Lee, J. R. & Taylor, J. H. *Science* **173**, 74–77 (1971).
- Ginsburg, A. P. *IEEE Transaction on Aerospace and Electronics* **9**, 625 (1973).
- Cavanagh, P. *Perception* **7**, 167–177 (1978).
- Schatz, B. R. *M.I.T. A.I. Memo* 426, (1977).
- Mayhew, J. E. W., Hall, G. & Frisby, J. P. *Perception* (submitted).
- Julesz, B. *Scient. Am.* **232**, 34–43 (1975).
- Marr, D. *Phil. Trans. R. Soc. B* **275**, 483–524 (1976).

Artificial menstrual cycles, behaviour and the role of androgens in female rhesus monkeys

THE sexual behaviour of higher primates has special relevance to the human situation. The rhesus monkey, for instance, has a true menstruation, and copulation occurs throughout the cycle. Levels of sexual activity, however, are generally higher in the peri-ovulatory period¹, and a similar mid-cycle maximum has been reported for certain human couples². Androgens increase the sexual behaviour of rhesus females^{3,4}, and they have been used to treat loss of libido in women⁵. These observations, together with the rise in plasma testosterone near mid-cycle, have led to suggestions that androgens may be the main libidinal hormone in the female⁶. We report here our experiments to test this hypothesis. We devised a daily dose schedule of oestradiol, testosterone and progesterone for ovariectomised rhesus monkeys that duplicated the changing plasma levels during normal menstrual cycles. When males were paired with females so treated, sexual activity showed a pattern similar to that during the normal cycle, and the omission of testosterone was without any detectable effect on this pattern.

Oestrogens given in physiological amounts to ovariectomised female rhesus monkeys increase the sexual behaviour of the pair, whereas injections of progestins antagonise the effects of oestrogen⁷. Testosterone given in pharmacological amounts to ovariectomised females both increases female sexual invitational behaviour⁸ and increases their lever-pressing for access to a male partner in an operant situation⁹. This has raised the question of whether the physiological increase in testosterone,

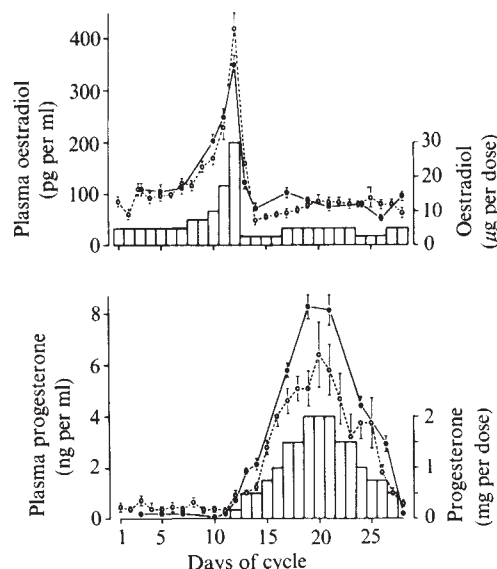


Fig. 1 Doses of oestradiol benzoate (upper histograms) and progesterone (lower histograms) given daily by subcutaneous injection to eight ovariectomised rhesus monkeys to imitate the changing plasma levels during normal cycles. Mean levels of oestradiol (upper graph) and progesterone (lower graph) in plasma samples drawn 8 h after injections during 24 artificial menstrual cycles (●) are compared with mean plasma levels for five intact females during 33 normal menstrual cycles (○). Each point represents the mean of 7–32 values. Vertical bars given standard errors of the means.

rather than the increase in oestrogen levels, near the middle part of the menstrual cycle may be primarily responsible for the menstrual rhythms in sexual behaviour.

Eight female and eight male rhesus monkeys were used in this study. The females were ovariectomised 3 months before the start of the experiment. Doses of oestradiol benzoate, progesterone and testosterone propionate in 0.2 ml sesame oil were administered subcutaneously at 08.00 h each morning according to the 28-d dose schedule shown in Figs 1 and 2. Females so treated had a normal menstruation every 28 d, and we have called these cycles 'artificial menstrual cycles'. Blood samples were collected at 16.00 h three or five times per week, and were analysed for oestradiol, progesterone and testosterone simultaneously using a radioimmunoassay involving chromatography on Sephadex LH20. Changes in the plasma levels of these steroids in 24 artificial menstrual cycles were closely similar to the levels in 33 normal cycles (Figs 1 and 2). Mean values were within the inter-individual variation for this species. When testosterone propionate was omitted from the daily dose schedule (eight cycles), plasma testosterone levels were steady throughout (adrenal secretion) and were similar to the levels in females receiving control injections of vehicle only (Fig. 2b). Thus, the administration of physiological doses of oestradiol and progesterone was without any detectable effect on the secretion of testosterone by the adrenals.

To determine the effect of these dose schedule on sexual behaviour, each female was paired alternately with one of two male partners (groups 1 and 2). Behaviour tests of 1-h duration were observed from behind one-way vision mirrors, and scores were entered directly into a computer through terminals in the observation booths. The number of ejaculations per test is used here as a measure of the males' sexual activity. To measure the females' motivation for propinquity with males, four of the females (eight pairs) were tested in an operant situation that permitted them to press a lever 250 times to gain access to a male partner. The females were allowed 30 min to perform this task and a 1-h behaviour test followed, either immediately, if access was attained, or later the same day, if it was not⁹.

When males were paired with females with artificial menstrual cycles, sexual activity showed a very consistent pattern similar to

that of the normal cycle. Both ejaculatory and operant performance were higher during cycles in which oestradiol, progesterone and testosterone were administered than during cycles with vehicle injections only, and in the latter cycles all signs of a 28-d behavioural rhythm were abolished. The mean frequency of ejaculations increased from 0.39 ± 0.05 per test ($n = 145$) in control cycles to 0.97 ± 0.04 per test ($n = 479$) in artificial cycles and, in operant behaviour tests, mean access time (latency to access) decreased from 909.5 ± 73.9 s ($n = 65$) in control cycles to 617.4 ± 30.4 s ($n = 239$) in artificial cycles. The omission of testosterone propionate from the dose schedule during eight cycles had no significant effect (analysis of variance) on either the mean frequency of ejaculations (0.94 ± 0.05 per test, $n = 156$) or mean access time (573.1 ± 55.1 s, $n = 80$). Comparable data collected from intact females (9 pairs, 33 cycles)¹ showed a mean frequency of 1.31 ± 0.04 ejaculations per test ($n = 447$), and a mean access time of 515.6 ± 23.6 s ($n = 443$).

Figure 3 shows changes in ejaculations, access times, female sexual invitations (presentations, hand-reaches, head-ducks and head-bobs) and in male success ratios (percentage of male mounting attempts not refused by the female) during these artificial cycles. Cyclic patterns of ejaculation occurred in all 24 cycles when all three steroids were administered and in all eight cycles when testosterone was omitted; there were no significant differences (analysis of variance) between cycles with and without testosterone. Access times were shortest in the week of highest oestradiol dosage in both types of artificial cycle. There were no consistent changes either in female proceptivity as measured by sexual invitations or in receptivity as measured by male success ratios, and there was clearly no evidence for reduced numbers of invitations or decreased male success ratios in cycles without testosterone.

These results are consistent with those of others¹⁰ in suggesting that the mid-cycle increase in plasma testosterone during normal menstrual cycles is of ovarian rather than of adrenal origin. Although functional adrenal glands (a major source of androgen in females) are necessary for the expression of many patterns of behaviour, including those of sexual behaviour¹¹, it seems unlikely that the physiological mid-cycle surge in testosterone plays any significant part in the increase in female sexual motivation at this time. The changes in behaviour observed in

Fig. 2 *a*, Daily doses of testosterone propionate given by subcutaneous injection (histograms) used in artificial menstrual cycles, and mean plasma levels of testosterone in 24 artificial cycles (●) compared with 33 normal cycles (○). *b*, Mean plasma levels of testosterone in eight artificial cycles (eight females) during which the schedule of oestradiol and progesterone was continued but in which testosterone propionate was omitted. The horizontal interrupted line represents the mean testosterone level (127.6 ± 25.7 pg per ml, $n = 45$) during control cycles when vehicle only was injected. Vertical bars give the s.e.m.

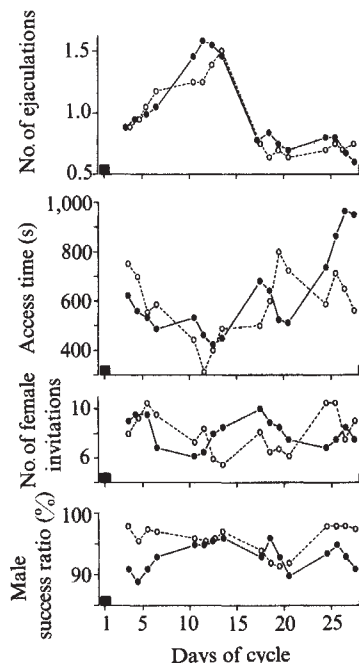
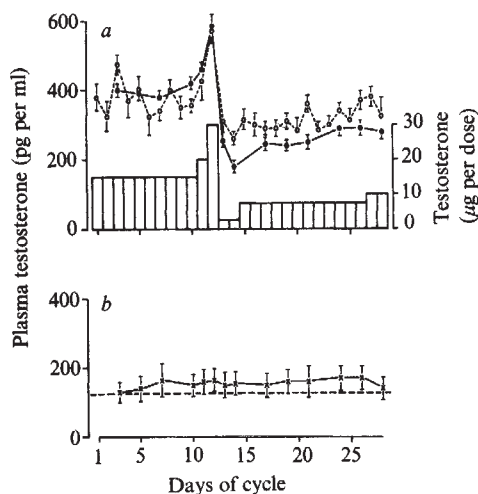


Fig. 3 Changes (means per test) in ejaculations, access times, female invitations and male success ratios during artificial menstrual cycles with oestradiol, progesterone and testosterone (●) and with oestradiol and progesterone alone (○) (2-d running means). Each point represents the mean of 8–48 tests. Black squares indicate menstruation.

the earlier studies cited were produced by doses (0.5–1.0 mg testosterone propionate) that were 15–400 times those used here (2.5–30.0 µg). These pharmacological doses produce plasma levels in the female rhesus monkey that are well within the normal range for males (4.0–25.0 ng per ml) rather than within the physiological range for females (100–800 pg per ml) that were obtained here. On the other hand, these studies show that normal behaviour cycles result from artificial cycles of physiological doses of oestradiol and progesterone. These findings support the view that the mid-cycle increase in oestradiol is sufficient to account for the increase in sexual activity during the peri-ovulatory period of the normal menstrual cycle.

This work was supported by grant no. MH19506 from the National Institute of Mental Health, and general support was received from the Georgia Department of Human Resources.

R. P. MICHAEL
M. C. RICHTER
J. A. CAIN
D. ZUMPE
R. W. BONSALE

Department of Psychiatry,
Emory University School of Medicine,
Atlanta, Georgia 30322
and
Georgia Mental Health Institute,
1256 Briarcliff Road, N.E.,
Atlanta, Georgia 30306

Received 2 May; accepted 7 August 1978.

1. Michael, R. P. & Bonsall, R. W. *Nature* **265**, 463–465 (1977).
2. Udry, J. R. & Morris, N. M. *J. Reprod. Fert.* **51**, 419–425 (1977).
3. Herbert, J. & Trimble, M. R. *Nature* **216**, 165–166 (1967).
4. Wallen, K. & Goy, R. W. *Horm. Behav.* **9**, 228–248 (1977).
5. Greenblatt, R. B., Mortara, F. & Torpin, R. *Am. J. Obstet. Gynec.* **44**, 658–663 (1942).
6. Money, J. in *Human Sexuality in Four Perspectives*, Vol. 3 (ed. Beach, F. A.) 62–86 (Johns Hopkins University Press, Baltimore, 1977).
7. Michael, R. P., Saayman, G. S. & Zumpe, D. *J. Endocr.* **41**, 421–431 (1968).
8. Michael, R. P. & Zumpe, D. *Anim. Behav.* **25**, 936–944 (1977).
9. Michael, R. P., Zumpe, D., Keverne, E. B. & Bonsall, R. W. *Recent Prog. Horm. Res.* **28**, 665–706 (1972).
10. Kim, M. H., Hosseinian, A. H. & Dupon, C. *J. clin. Endocr. Metab.* **39**, 706–712 (1974).
11. Everitt, B. J., Herbert, J. & Hamer, J. D. *Physiol. Behav.* **8**, 409–415 (1972).

Elaborate CNS cooling structures in large dinosaurs

DURING sustained activity, or when exposed to high ambient temperatures, terrestrial animals often experience periods of core hyperthermia. The central nervous system (CNS) is very sensitive to elevated temperatures¹⁻³, and consequently, both bradymetabolic and tachymetabolic⁴ terrestrial vertebrates have evolved physiological mechanisms which effect localised cooling of the brain, and thereby reduce any thermal impairment of CNS functioning. In modern reptiles this temperature gradient is produced by evaporative cooling from the buccal cavity and upper respiratory tract^{5,6}, conducting heat from the brain through the floor of the cranium⁷. Mammals dissipate heat, by evaporation, from the nasal mucosa to the air flowing through the nasal passages. The cooled venous blood draining from this highly vascularised mucosa flows into the cavernous sinus, where counter-current heat exchange with the carotid arteries, elaborated into a rete in many forms, results in brain cooling^{8,9} (Fig. 1a). Dinosaurs would have experienced similar thermal problems to those of modern vertebrates, and these would have been particularly acute in the larger forms whose low surface area to volume ratio would have restricted dissipation of the enormous amounts of heat generated by the skeletal muscles during activity. It is therefore proposed that they required and possessed comparable physiological mechanisms to protect the brain during core hyperthermia.

In the lambeosaurine hadrosaurs, air entering the nostrils passed through extensive passages in a crest formed by the premaxillae and nasal bones before descending the choanal tube to the throat (Fig. 2a). Early workers associated the function of these crests with a semi-aquatic mode of life. The theories proposed included air storage chambers or traps to prevent water entering the lungs¹⁰, but evidence now indicates that hadrosaurs were predominantly terrestrial^{11,12}. Ostrom^{13,14} showed the early theories to be untenable and concluded that the function of the crests was to increase the area of mucosal

Fig. 1 a, Schematic diagram of the blood vessels involved in brain cooling in the sheep (after Hayward and Baker⁹). b, Proposed mechanism of brain cooling in *Edmontosaurus*. Cooled venous blood from the nasal capsule drains into the enlarged orbital sinuses where it conducts heat from the brain.

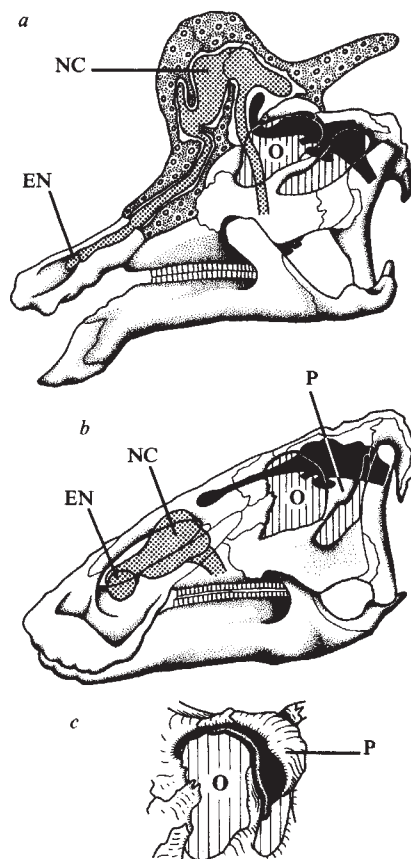
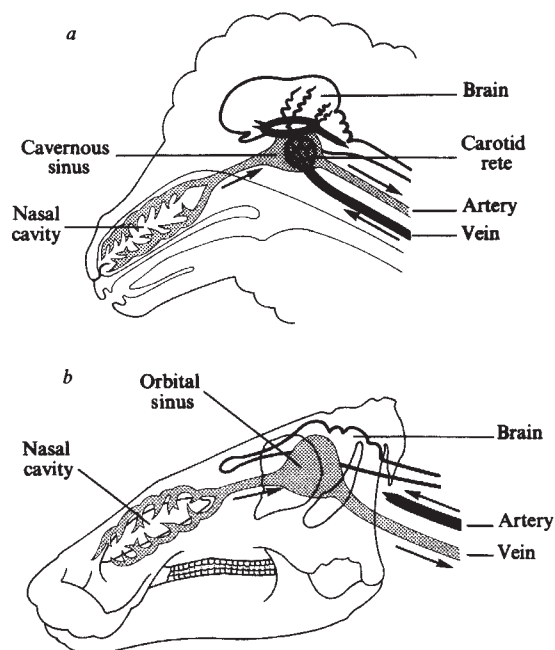


Fig. 2 a, Section through the narial crest of the hadrosaur *Lambeosaurus* illustrating the expanded nasal passages and their position relative to the brain (black). EN, external naris; NC, nasal capsule; O, orbit (after Ostrom¹⁴). b, Skull of the hadrosaurine *Edmontosaurus* showing probable position of the cartilaginous nasal capsule (after Ostrom¹⁴). c, Enlargement of the orbit and postorbital pouches (P) of *Edmontosaurus* (after Ostrom¹³).

epithelium lining the nasal passages, suggesting that this enlargement was for enhanced olfaction. However, it seems likely that the function was primarily thermoregulatory. Although probably too small to have an appreciable effect on core temperature, the enlarged mucosal area would have increased the rate of evaporative heat loss from the head and hence prevented excessive elevation of brain temperature. This idea is supported by the proximity of the passages to the cranium, which would have been cooled by direct conduction into the crest and by the action of the cooled venous blood draining from the mucosa. As in modern mammals, the rate of cooling may have been related to the regulation of air flow through the nasal passages, with panting occurring during heat stress¹⁵. This proposal is not incompatible with the suggestion^{16,17} that these crests, and those of the saurolophine hadrosaurs, also functioned in visual display.

The non-lambeosaurine hadrosaurs lacked hollow narial crests, but their external nares were conspicuously expanded and probably contained an enlarged cartilaginous nasal capsule¹⁴ (Fig. 2b), the lining of the channels within which would have formed an extensive evaporative surface. Unlike the lambeosaurine crests, the location of this surface indicates little influence on brain temperature by direct heat conduction; however, the venous blood draining from it would have cooled the cranium. Evidence for this is the presence, in some genera, of large pouch-like cavities in the postorbital bones (Fig. 2c). Ostrom¹³ suggested these resulted from enlargement of the orbital sinuses, venous structures which in modern reptiles collect blood from the surrounding areas of the head and empty into the internal jugular veins. If this interpretation is correct the cooled venous blood from the nasal capsule would have drained

into these enlarged sinuses, the proximity of which to the cranium would have resulted in brain cooling (Fig. 1b). Furthermore, counter-current heat exchange may have occurred between the jugular venous and carotid arterial blood, similar to that described in modern mammals⁸ and reptiles¹⁸.

Many sauropods greatly exceeded 10 tonnes in weight and were the largest known dinosaurs. The external nares of these animals were situated high on the skull and were usually very large, indicating a spacious nasal capsule (Fig. 3). This location of the nares was thought to have been an aquatic adaptation, allowing these animals to breathe while almost totally submerged¹⁰; but the nares of most aquatic reptiles are small, unlike those of sauropods. From the structure of the vertebral column and limbs, Bakker^{19,20} proposed that sauropods were terrestrial although he was unable to clarify the functional significance of the dorsal position of the nares. However, both the size and location of the nasal capsule are consistent with its having an important role in brain thermoregulation, its enlargement increasing the mucosal area available for evaporation and its position reducing the distance between this cooling surface and the cranium. The long neck of these dinosaurs increased the thermal isolation of the head from the body and extended the distance available for heat exchange between the jugular veins and carotid arteries. This explanation of narial structure strongly supports the evidence that sauropods were terrestrial¹⁹⁻²¹ and argues against the aquatic theory.

As in sauropods, hadrosaurs and large ornithiscian dinosaurs such as *Iguanodon*, the nasal region of the skull of the ceratopsians was greatly enlarged and again the nasal mucosa was probably the main site of heat dissipation. In modern ruminants²² the surface of the bony horn cores is covered with a rich vascular plexus which vasodilates during heat stress, resulting in increased blood flow through the horns. This blood, cooled by convection and radiation from the surface of the horns, drains into the cavernous sinus where it comes into contact with the carotid rete and cools the arterial blood entering the brain. Most ceratopsians possessed one or more horns with cores very like those of ruminants¹⁰. A network of channels on their surface indicates that they too were highly vascularised and therefore could have functioned comparably in brain thermoregulation. Vascularisation of the surface of the ceratopsian

parietal-squamosal neck frills, especially the marginal areas not overlain by jaw musculature²³, would also have permitted additional heat loss from the head.

In the carnosaurs, whose external nares are not as enlarged as those of other large dinosaurs, a major site of heat loss was probably, as in modern reptiles, the lining of their enormous buccal cavity. In some forms there was considerable flexibility between the various jaw elements, structural modifications thought to have facilitated the engulfing of large portions of prey¹⁰. This could indicate the presence of a fairly extensive gular area which would have also further extended the evaporative surface area. Additional cooling may have been achieved by active pumping or fluttering of this gular region, a method used by large varanid lizards²⁴ and birds such as the ostrich²⁵.

Some large dinosaurs evolved structures which can be interpreted as providing thermal protection for the extracranial CNS. The vertebrae of many forms contain deep excavations which probably contained an air sac system similar to that of birds²⁶. Ventilation of this system would have removed heat from the region around the spinal cord. Not surprisingly, the excavations are particularly extensive in the vertebrae of sauropods²¹, which, because of their size, would have been especially vulnerable to hyperthermia. *Stegosaurus* possessed additional radiative and convective structures consisting of a double row of bony plates, well permeated by blood vessels, running along the neck, back and tail²⁷. Although the size of these plates suggests they significantly influenced core temperature, the spinal cord would have been particularly well protected from elevated temperatures by the cooled blood draining from the plates directly above it. The largest plates are located in the sacral region where the spinal cord was apparently enlarged into a ganglion exceeding the brain in size. The vertebral neural spines of a few dinosaurs, such as *Spinosaurus* and *Ouranosaurus*, were extremely elongated and are thought to have supported a vascular sail-like structure which would similarly have permitted increased radiative and convection heat loss, so cooling the spinal region.

The described CNS cooling mechanisms do not constitute direct evidence that dinosaurs were either bradymetabolic or tachymetabolic animals, as such physiological mechanisms are found in the extant groups of both. However, the apparent very highly developed nature of these structures in dinosaurs does tend to support the idea that they may have been tachymetabolic, with a high metabolic level and resulting endogenous heat production, rather than bradymetabolic like modern reptiles. Unfortunately, the nervous and vascular systems are not preserved in fossils and it is only possible to observe the structure of those portions which leave an impression on the skeleton. This makes it impossible to establish the detailed physiological functioning of the proposed cooling mechanisms, but it seems reasonable to assume that they would have been as sophisticated as those of modern vertebrates, incorporating heat-exchanging retes and vascular shunts^{18,28} allowing control of blood flow to and from the cooling surfaces.

I thank Dr K Bowler for discussions, and Professor J. E. Heath and Dr T. S. Kemp for critical reading of the manuscript.

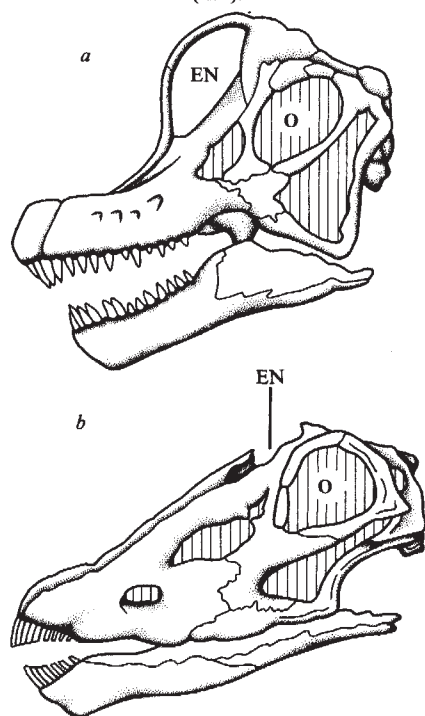
P. E. WHEELER

Department of Zoology,
University of Durham,
South Road, Durham, UK

Received 19 April; accepted 29 August 1978.

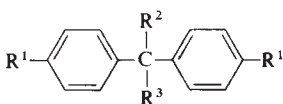
1. Precht, H., Christopherson, J., Hensel, H. & Larcher, W. *Temperature and Life* 656-658 (Springer, Berlin, 1973).
2. Becht, F. C. *Am. J. Physiol.* **22**, 456-476 (1908).
3. Webb, G. J. W. & Witten, G. J. *Comp. Biochem. Physiol.* **45A**, 829-832 (1973).
4. Bligh, J. & Johnson, K. G. *J. appl. Physiol.* **35**, 941-961 (1973).
5. Crawford, E. G., Palomeque, J. & Barber, B. J. *Comp. Biochem. Physiol.* **56A**, 161-163 (1977).
6. Webb, G. J. W. & Johnson, C. R. *Comp. Biochem. Physiol.* **43A**, 593-611 (1972).
7. Crawford, E. C. *Science* **177**, 431-433 (1972).
8. Baker, M. A. & Hayward, J. N. *J. Physiol., Lond.* **198**, 561-579 (1968).
9. Hayward, J. N. & Baker, M. A. *Brain Res.* **16**, 417-440 (1969).
10. Romer, A. S. *Vertebrate Paleontology* (University of Chicago Press, 1966).
11. Galton, P. M. *J. Paleontol.* **44**, 464-73 (1970).
12. Ostrom, J. H. *Am. J. Sci.* **262**, 975-997 (1964).
13. Ostrom, J. H. *Am. Mus. nat. Hist. Bull.* **122**(2), 33-186 (1961).

Fig. 3 Skulls of the sauropods *Brachiosaurus* (a) and *Diplodocus* (b), illustrating the size and dorsal location of the external nares (EN).



14. Ostrom, J. H. *Postilla* **62**, 1–29 (1962).
15. Young, B. A., Bligh, J. & Louw, G. J. *Thermal Biol.* **1**, 195–198 (1976).
16. Dodson, P. *Syst. Zool.* **24**, 37–54 (1975).
17. Hopson, J. A. *Paleobiology* **1**, 21–43 (1975).
18. Heath, J. E. *Physiol. Zool.* **39**, 30–35 (1966).
19. Bakker, R. T. *Discovery* (New Haven) **3**, 11–22 (1968).
20. Bakker, R. T. *Nature* **229**, 172–174 (1971).
21. Coombs, W. P. *Paleogeogr., Palaeoclimatol., Palaeocol.* **17**, 1–33 (1975).
22. Taylor, C. R. *Physiol. Zool.* **39**, 127–139 (1966).
23. Ostrom, J. H. *Evolution* **20**, 290–308 (1966).
24. Johnson, C. R. *Comp. Biochem. Physiol.* **43A**, 1025–1029 (1972).
25. Schmidt-Nielsen, K. et al. *Condor* **71**, 341–352 (1969).
26. Bakker, R. T. *Evolution* **25**, 636–658 (1971).
27. Farlow, J. O., Thompson, C. V., Rosner, D. E. *Science* **192**, 1123–1125 (1976).
28. Magilton, J. H. & Swift, C. S. *J. appl. Physiol.* **27**, 18–20 (1969).

Table 1 Structure of the DDT analogues tested

			
Compound	R ₁	R ₂	R ₃
<i>p-p'</i> -DDT	Cl	H	CCl ₃
TDE	Cl	H	CHCl ₂
Perthane	C ₂ H ₅	H	CCl ₂
Methoxychlor	CH ₃ O	H	CCl ₃
Ethoxychlor	C ₂ H ₅ O	H	CCl ₃
Chlorphenethol	Cl	OH	CH ₃
Dicofol	Cl	OH	CCl ₃
FDMC	Cl	OH	CF ₃
Chlorobenzylate	Cl	OH	COOC ₂ H ₅
Chloropropylate	Cl	OH	COOCH(CH ₃) ₂

Unusual response of DDT-resistant houseflies to carbinol analogues of DDT

THE toxic action of DDT is thought to be caused by complex formation between the insecticide and nerve protein, a mechanism consistent with the well known phenomenon of negative correlation between insect mortality and temperature¹. We present here evidence which indicates that some carbinols which are close analogues of DDT may have a different mode of action. These compounds elicit very different symptoms of poisoning, and none of the major mechanisms of resistance to DDT in the housefly confers significant resistance to these compounds.

Several analogues of DDT (Table 1) were tested against strains of houseflies (*Musca domestica* L.) with known genetically isolated mechanisms of resistance to DDT. These strains were: Beiruth (heterozygous for *DDTase*), 2778 (*DDT md*, monooxygenase, sesamex-suppressible resistance), 538 ge (*kdr*-resistance)² and Super-*kdr* (*kdr*-like resistance). This last mechanism, not yet fully characterised, was recently isolated from the multi-resistant 153y3 strain of houseflies by the chromosome isolation technique³. Super-*kdr* resembles *kdr* in many respects; it is controlled by one or more incompletely recessive factor(s) on chromosome 3, confers resistance not only to DDT and its analogues but also to the pyrethroids, and is unaffected by pretreatment with FDMC (bis(*p*-chlorophenyl)trifluoromethyl carbinol; 1 µg per fly), piperonyl butoxide (2 µg per fly) or the two additives applied together. Also, it crosses freely with *bwb* (brown body, a recessive mutant marker for chromosome 3), confers no detectable resistance to organophosphorus or carbamate insecticides and differs from *kdr* only in its greater or much greater resistance towards DDT-like compounds and the pyrethroids (Table 2).

Three- to five-day-old female houseflies of the four strains were treated at room temperature with the compounds dissolved in acetone, and were held either for 48 h at 20 °C or 3–4 d at 15, 20 and 28 °C, and were checked daily for numbers killed. Of the carbinols tested (Table 1) only dicofol and FDMC had relatively strong insecticidal activity.

DDT and four of its analogues (Table 1) produced typical symptoms of DDT poisoning in the susceptible and three of the resistant flies. These symptoms are increased activity, tremors,

incoordination and tetany. Toxic symptoms were delayed in the resistant flies, particularly in the Super-*kdr* strain, in which they became visible at least several hours after treatment with DDT. The carbinols produced different symptoms which were the same for all the strains tested. The initial sign of poisoning, increased locomotor activity, was first apparent several hours after treatment even at the strongest dose and the highest holding temperature. The insects then exhibited transient collapse with a rapid return to normal activity. The periods of prostration, initially infrequent and very brief, increased in duration and frequency until the insects were no longer able to right themselves or walk. Tremors, incoordination of the limbs and increased wing-beat observed in DDT-poisoned flies were absent or rare in the carbinol-treated insects. Paralysis was neither induced nor reversed by altering the holding temperature, but the insects were killed faster at the higher temperature. At death (end-point), the relative toxicities of the carbinols varied relatively little for the three holding temperatures (Table 3), but those for DDT and its analogues were negatively correlated with temperature (Table 3) in all strains except Super-*kdr*, which had a positive temperature coefficient for DDT (Table 3). This is almost certainly due to the very slow action of this compound against those flies at 20 and 15 °C, temperatures at which death is delayed well beyond 4 d. Perthane, which acts much faster against Super-*kdr* flies than DDT, gave a negative temperature coefficient.

The other major difference with the carbinols was the slight or negligible resistance of the DDT-resistant strains to these compounds (Table 3). This lack of effect by major known mechanisms of resistance to DDT in the housefly was unexpected, particularly so for *kdr*, which is believed to alter either access of insecticide to the site of action, or the site of action itself, presumably a nerve-membrane receptor for DDT-type insecticides¹. (*kdr* confers resistance to DDT analogues that

Table 2 Cross resistance of strain 538 ge and Super-*kdr* to several insecticides 48 h after treatment at 20 °C

Insecticide	LD ₅₀ µg per fly $\pm$ s.e. (resistance factor)								
	Susceptible			538 ge			Super- <i>kdr</i>		
Malathion	0.37	1.10	(1.0)	0.40	1.08	(1.1)	0.36	1.09	(0.97)
Parathion	0.016	1.08	(1.0)	0.016	1.06	(1.0)	0.018	1.09	(1.1)
Bioresmethrin	0.010	1.10	(1.0)	0.045	1.08	(4.5)	0.31	1.10	(31)
Cismethrin*	0.0040	1.14	(1.0)	0.072	1.10	(18)	1.5	1.11	(375)
Decamethrin†	0.00049	1.09	(1.0)	0.0069	1.10	(14)	0.30	1.10	(612)
DDT	0.022	1.07	(1.0)	0.48	1.09	(22)	15.0	1.15	(682)

* 5-Benzyl-3-furylmethyl(IR,*cis*)-chrysanthemate.

† S-(α -cyano-3-phenoxybenzyl)(IR,*cis*)-[3-(2,2-dibromovinyl)]-2,2-dimethylcyclopropanecarboxylate.

Table 3 Response of several strains of houseflies to some DDT analogues at three holding temperatures

		LD_{50} μg per fly $\times$ s.e.														
Housefly strain resistance mechanism		Susceptible (none)			Beiruth DDTase			2778 monooxygenase			538 ge <i>kdr</i>		Super- <i>kdr</i> <i>kdr</i> -like			
Compound	Holding temperature															
DDT	15	0.0075	1.07	(1.0)*		**		0.15	1.09	(1.0)	0.68	1.09	(1.0)	36.0	1.32	(1.0)
	20	0.020	1.09	(2.7)		**		0.35	1.11	(2.3)	0.84	1.09	(1.2)	15.8	1.15	(0.44)
	28	0.041	1.10	(5.5)		**		1.12	1.11	(7.4)	1.23	1.11	(1.8)	9.0	1.20	(0.25)
Perthane	15	0.048	1.08	(1.0)	0.21	1.13	(1.0)	0.21	1.11	(1.0)	0.89	1.08	(1.0)	2.45	1.09	(1.0)
	20	0.11	1.08	(2.3)	0.34	1.12	(1.6)	0.80	1.14	(3.8)	1.64	1.09	(1.8)	5.91	1.13	(2.4)
	28	0.25	1.10	(5.2)	—	—	—	1.69	1.09	(8.0)	2.05	1.08	(2.3)	8.38	1.11	(3.4)
TDE	15	0.16	1.07	(1.0)		**		0.26	1.09	(1.0)	1.25	1.08	(1.0)			***
	20	—	—	—		**		0.63	1.08	(2.4)	2.64	1.11	(2.1)			***
	28	0.44	1.08	(2.8)		**		—	—	—	4.12	1.13	(3.3)			***
Ethoxychlor	15	0.029	1.10	(1.0)		**		0.20	1.08	(1.0)	1.59	1.12	(1.0)			***
	20	0.064	1.07	(2.2)		**		1.85	1.10	(9.3)	13.2	1.31	(8.3)			***
	28	0.095	1.10	(3.3)		**		—	—	—	***					
Methoxychlor	15	0.073	1.06	(1.0)		**		0.86	1.10	(1.0)	1.80	1.09	(1.0)			***
	20	0.076	1.10	(1.0)		**		2.03	1.09	(2.4)	3.95	1.14	(2.2)			***
	28	0.12	1.11	(1.6)		**		—	—	—	43.8	1.35	(24.3)			***
FDMC	15	1.85	1.08	(1.0)	2.09	1.06	(1.0)	2.49	1.08	(1.0)	1.17	1.07	(1.0)	2.92	1.08	(1.0)
	20	2.27	1.09	(1.2)	3.23	1.07	(1.5)	2.68	1.09	(1.1)	2.16	1.09	(1.8)	2.81	1.08	(0.96)
	28	1.65	1.17	(0.89)	2.51	1.09	(1.2)	3.85	1.16	(1.5)	2.26	1.09	(1.9)	2.37	1.11	(0.81)
Dicofol	15	1.07	1.25	(1.0)	2.12	1.10	(1.0)	—	—	—	1.93	1.08	(1.0)	3.79	1.09	(1.0)
	20	1.39	1.14	(1.3)	2.86	1.12	(1.3)	2.6	1.11	—	2.02	1.08	(1.0)	2.69	1.09	(0.71)
	28	1.41	1.15	(1.3)	2.11	1.14	(1.0)	—	—	—	2.47	1.13	(1.28)	2.36	1.09	(0.62)

* Numbers in parentheses indicate temperature coefficient, LD_{50} at given temperature/ LD_{50} at 15 °C.

** Strain heterogeneous for DDTase. Hence, LD_{50} s for insects with strong DDTase not measurable.

*** LD_{50} s not obtained with doses exceeding 40 μg insecticide per fly.

cannot be dehydrochlorinated and is responsible for low nerve sensitivity to DDT directly applied to the thoracic ganglion⁴). Dicofol and FDMC are therefore not only potent inhibitors of DDT dehydrochlorinase but also affect a site of action different from that primarily affected by DDT. This site is unknown, but several enzymes are known to be affected by DDT⁵, particularly the oligomycin-sensitive Mg^{2+} ATPase⁶. Chang and Cutkomp⁷ have demonstrated that inhibition of this enzyme in cockroach muscle by DDT, TDE, methoxychlor and DDE has a negative temperature coefficient, whereas the activity of dicofol and chlorphenethol increases with temperature.

It is interesting to speculate whether the major target of attack by carbinols in acarines is similar to that in houseflies, as acarines are very susceptible to most of the carbinols tested here but are unaffected by DDT and many of its analogues. Also, it remains to be seen whether the slow toxic action described here is confined only to the carbinol analogues of DDT, as the screening procedures used so far for bioassays on analogues of DDT were not designed to detect this type of action.

I thank Dr I. J. Graham-Bryce for discussions, Ciba-Geigy and Murphy Chemical for some of the carbinols, and Drs J. Keiding and A. W. Farnham for two of the resistant strains of houseflies.

R. M. SAWICKI

Department of Insecticides and Fungicides,
Rothamsted Experimental Station,
Harpenden, Herts, UK

Received 27 July; accepted 4 September 1978.

- Holan, G. *Nature* **221**, 1025–1029 (1969).
- Farnham, A. W. *Pestic. Sci.* **8**, 631–636 (1977).
- Tsukamoto, M. *Botyu-Kagaku* **29**, 51–59 (1964).
- Tsukamoto, M., Nara-hashi, T. & Yamasaki, T. *Botyu-Kagaku* **30**, 128–132 (1965).
- Perry, A. S. *Misc. Publ. ent. Soc. Am.* **2**, 119–137 (1960–61).
- Cutkomp, L. K., Yap, H. H., Vea, E. A. & Koch, R. B. *Life Sci.* **10**, 1201–1209 (1971).
- Cheng, E. Y. & Cutkomp, L. K. *Pestic. Biochem. Physiol.* **7**, 360–366 (1977).

Chloroquine enhances Epstein-Barr virus expression

AFRICAN Burkitt's lymphoma (BL) is a malignant tumour of children most common in areas having holoendemic malaria^{1,2}. It occurs with a particularly high frequency in certain areas of Africa³, with sporadic cases in other areas including Europe and North America^{4–7}. Malaria has been shown to have immunosuppressive effects^{8,9}, and chronic malaria, in particular, causes repeated stimulation of the immune response to infections with plasmodia. It has also been proposed that the immunosuppressive action of malarial drugs may make the immune system more susceptible to cancer and may contribute to BL development. Chloroquine is widely used as an anti-malarial agent and also in the therapy of connective tissue diseases^{10,11}. It is thought to act by forming molecular complexes with plasmodial DNA, causing inhibition of plasmodial DNA synthesis^{12–14}. It has been suggested that chloroquine may play a part in the development of BL¹⁵. Another factor thought to be involved in the aetiology of the African form of BL is the Epstein-Barr virus. Virtually all BLs are EBV genome-carrying tumours and patients have high titres of antibody against EBV^{16,17}. The most widely proposed hypothesis is that BL is due to EBV-induced malignant transformation of thymus-independent B lymphocytes^{17,18}. It has been shown *in vitro* that EBV can induce transformation of resting human and simian lymphocytes which become established as permanent lines^{19,20}. Almost all the cells in these lines contain EBNA (the EBV-induced nuclear antigen) and carry multiple copies of EBV DNA^{21,22}. It has also been shown that EBV can induce lymphomas in experimentally infected monkeys²³. However, the failure of EBV infection to have serious adverse effects in the great majority of humans indicates that interaction with some

other factor, such as malaria or antimalarial drugs, must be involved in BL pathogenesis. We report here studies on the effect of chloroquine on the expression of EBV antigens. The data obtained indicate that chloroquine can activate the EBV cycle and the expression of EBV antigens.

EBV-infected cells express one or more of the EBV-induced antigens, which are readily detectable by immunofluorescence with the use of human sera containing anti-EBV antibodies. These antigens include EBNA, early antigen (EA), virus capsid antigen (VCA) and membrane antigen (MA). The cell line used here was Raji²⁴, which is EBNA-positive, EA-negative and VCA-negative. The cells were grown in medium RPMI 1640 supplemented with 10% heat-inactivated fetal bovine serum, penicillin (100 U ml⁻¹) and streptomycin (100 µg ml⁻¹). Raji cells were superinfected with P3HR-1 strain of EBV²⁵ for the induction of EA and VCA. Drug treatment was carried out as follows: 2 × 10⁶ lymphoid cells were sedimented by low-speed centrifugation and the cells were washed three times with phosphate-buffered saline (PBS). The pellet was dispersed and incubated at 37°C in 5 ml of medium containing chloroquine (diphosphate salt, Sigma) at concentrations of 0, 0.1, 1.0, 5.0 µg ml⁻¹ (0, 1.9 × 10⁻⁷, 1.9 × 10⁻⁶, 9.5 × 10⁻⁶ M). Six tubes were prepared for each treatment and the viability of the cells (as determined by the trypan blue exclusion test) was estimated after 20–24 h. Only cell preparations which contained less than 5% nonviable cells were used for EBV infections. 200,000 cells from each of these preparations were pelleted by low-speed centrifugation. The supernatant was discarded and 0.1 ml virus suspension was added per pellet, except to the control. The virus-cell mixture was shaken gently, 1 ml of medium containing the appropriate amount of chloroquine was added and the cultures incubated for 48 h at 37°C. Cells were then collected, washed three times in PBS and resuspended in it. Four smears were prepared from each culture, dried in air at room temperature, and fixed in a 1:1 mixture of acetone-methanol at -20°C for 3 min. The smears were kept at -20°C until used for staining by the standard immunofluorescence methods described in Table 1.

The data on the effect of chloroquine on superinfection of Raji cells with P3HR-1 EBV are shown in Table 1. Chloroquine treatment caused a significant increase in EA-positive cells at a concentration as low as 0.1 µg ml⁻¹ (1.9 × 10⁻⁷ M). In this regard, it is interesting that plasma concentrations of chloroquine in man during malaria chemotherapy range from 3 × 10⁻⁷ to 10⁻⁶ M (0.16 to 0.53 µg ml⁻¹)²⁶. Above 0.1 µg ml⁻¹ chloroquine, there was no further increase in the number of EA-positive cells. There was no difference in VCA-positive cells in the presence of chloroquine.

In preliminary experiments we found that chloroquine had no effect on EA and VCA in EBV genome-carrying cells not

superinfected by EBV, thus indicating that unlike iododeoxyuridine²², chloroquine does not induce antigen expression in non-exogenously infected cells.

The results presented here clearly indicate that chloroquine increases EA expression in superinfected cells. This suggests that chloroquine may be acting either directly on the exogenous EBV genome or indirectly by activating cellular mechanisms essential for the expression of the exogenous EBV genome.

It is well known that BL patients have high antibody titres against EBV antigens^{16,17}. In the present study, chloroquine gave rise to an increased expression of EA, an EBV antigen, which *in vivo* would be expected to provoke an increased antibody reaction. It would therefore be useful to compare levels of antibody titres in BL patients who received chloroquine therapy with control BL patients who were never treated with this drug.

Although the exact mechanism by which chloroquine treatment affects viral expression is unknown, one possibility is that chloroquine alters the relationship between DNA and cellular controls and this hypothesis is now being investigated. Chloroquine can also form molecular complexes with DNA and inhibit DNA-dependent nucleic acid polymerase reactions^{12–14}.

In conclusion, the present study suggests that chloroquine may play an important part in the development of African Burkitt's lymphoma by enhancing the expression of EBV, an oncogenic human herpesvirus. This does not exclude a further contribution resulting from the immunosuppressive action of chloroquine.

J. Menezes is a scholar of the MRC of Canada. These studies were supported by the NCI of Canada.

R. A. KARMALI*
D. F. HORROBIN

Clinical Research Institute, 110 Pine Avenue West,
Montreal

J. MENEZES
P. PATEL
J. MUSTO

Department of Microbiology and Immunology,
Pediatric Research Center, University of Montreal,
Ste-Justine Hospital, 3175 Ste-Catherine Rd,
Montreal, Ontario, Canada

Received 27 April; accepted 7 August 1978.

*Present address: Department of Pathology, School of Medicine, East Carolina University, Greenville, North Carolina 27884, USA.

Table 1 Effect of chloroquine on % EA-positive cells

	Raji	Raji + EBV	Raji + EBV + chloroquine (µg ml ⁻¹)		
			0.1	1.0	5.0
EA	0	6.5 ± 0.95	19.0 ± 1.61 (P < 0.001)	16.8 ± 1.83 (P < 0.001)	18.3 ± 1.97 (P < 0.001)

The tests for EA and VCA were carried out by indirect immunofluorescence. The smears were incubated at room temperature for 45 min with anti-EA serum LM (with EBNA-positive, EA-positive and VCA-positive antibody composition) diluted 1:20, and with anti-VCA reference serum JM diluted 1:10. These smears were washed three times with PBS and incubated for 45 min with anti-IgG conjugate (Highland) diluted 1:30. The smears were then washed again three times with PBS, counterstained with Evans blue for 10 min, washed three times with PBS and mounted on a slide in a 1:1 mixture of glycerol-PBS. Raji cells were used as control for EA and B95-8 cells¹⁹ as control for VCA. EA- and VCA-positive cells were estimated as a % of approximately 300 cells. Each concentration was tested in six cultures and the results show the mean ± s.e.m.

- Burkitt, D. P. *Int. Rev. exp. Path.* **2**, 67–138 (1963).
- Klein, G. in *The Herpes Viruses* (ed. Kaplan, A.) 521–555 (Academic, New York, 1973).
- Burkitt, D. Br. J. Surg. **46**, 218–223 (1958); *Nature* **194**, 232–234 (1962).
- Pagano, J. S., Huang, C. H. & Levine, P. *New Engl. J. Med.* **289**, 1395–1399 (1973).
- Levine, P. H. et al. *Ann. intern. Med.* **83**, 31–36 (1975).
- Ziegler, J. L. *New Engl. J. Med.* **297**, 75–80 (1977).
- Judson, S. C. & Henle, G. *New Engl. J. Med.* **297**, 464–468 (1977).
- Salaman, M. H., Wedderburn, N. & Bruce-Chwatt, L. J. *J. gen. Microbiol.* **59**, 383 (1969).
- Gazayerli, M. *Bull. Int. Acad. Path.* **4**, 37 (1963).
- Goodman, L. S. & Gilman, A. in *The Pharmacological basis of Therapeutics*, 5th edn, 1049–1053 (Macmillan, London, 1975).
- Coatney, G. R. *Am. J. trop. Med. Hyg.* **12**, 121–128 (1963).
- Parker, F. S. & Irvin, J. L. *J. biol. Chem.* **199**, 897 (1952).
- Cohen, S. N. & Yielding, K. L. *Proc. natn. Acad. Sci. U.S.A.* **54**, 521–527 (1965).
- Karmali, R., Schiller, P. W. & Horrobin, D. F. *Prostaglandins* **12**, 463–464 (1976).
- Sadoff, L. *Lancet* **ii**, 1262–1263 (1973).
- Henle, G. et al. *J. natn. Cancer Inst.* **43**, 1147–1157 (1969).
- Klein, G. *New Engl. J. Med.* **293**, 1353–1357, 1975.
- Olweny, C. L. M. et al. *J. natn. Cancer Inst.* **58**, 1191–1196 (1977).
- Pope, J. H., Horne, M. K. & Scott, W. *Int. J. Cancer* **3**, 857–866 (1968).
- Miller, G., Shope, T., Lisco, H., Stit, D. & Lipman, M. *Proc. natn. Acad. Sci. U.S.A.* **69**, 383–387 (1972).
- Lindahl, T. et al. *Int. J. Cancer* **13**, 764–772 (1974).
- Menezes, J., Leibold, W. & Klein, G. *Expl Cell Res.* **92**, 478–484 (1975).
- Shope, T., Dechario, D. & Miller, G. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2487–2491 (1973).
- Pulvertaft, R. J. V. *J. clin. Path.* **18**, 261–273 (1965).
- Hinuma, Y. & Grace, J. T. *Proc. Soc. exp. Biol. Med.* **124**, 107–111 (1967).
- Covell, G., Coatney, G. R., Field, J. W. & Singh, J. *Monograph Ser. W.H.O.* **27** (1955).

Asbestos-mediated membrane uptake of benzo[a]pyrene observed by fluorescence spectroscopy

EPIDEMIOLOGICAL evidence indicates that particulate matter, such as asbestos, can increase the health risk from exposure to chemical carcinogens. For example, asbestos insulation workers show an eightfold increased disposition towards lung cancer; non-smokers had only a slightly increased risk of lung cancer, whereas the smokers had a 92-fold increased risk^{1,2}. These data indicate that 90% of the asbestos related cancers are the result of the co-carcinogenic effects of cigarette smoking and the inhalation of asbestos fibres. We have now examined one possible mechanism of co-carcinogenesis—the ability of particulate matter to increase the rate of cellular uptake of benzo[a]pyrene (BP). We used fluorescence spectroscopy to investigate the ability of two particulates, silica and asbestos, to adsorb BP, and to deliver the carcinogen to model membranes. The surface area of the asbestos sample was 60-fold less than that of our silica sample. Surprisingly the asbestos sample showed a greater ability to adsorb BP in a monomeric state than silica. Most importantly BP which is adsorbed to asbestos is more rapidly transported into membranes than BP which is adsorbed to silica. Thus increased cellular exposure to carcinogens, resulting from the presence of particulates which can adsorb and deliver these carcinogens to cells, may be involved in the process of co-carcinogenesis.

Figure 1 shows fluorescence emission spectra of BP in dilute benzene solution, and of a suspension of BP crystals in aqueous buffer. A structured emission is observed from BP dissolved in benzene, whereas the suspension of BP crystals yields a broad structureless emission at longer wavelengths. This latter emission results from formation of an excited state charge-transfer complex between two or more adjacent BP molecules³ and is called the excimer emission. The fluorescence emission spectra may therefore be used to reveal the proximity of BP molecules to one another. Also shown in Fig. 1 is the fluorescence emission spectrum of BP which is partitioned into lipid bilayers after uptake from asbestos fibres. Only the structured emission is observed which is indicative of widely dispersed BP molecules.

Fig. 1 Fluorescence emission spectra of benzo[a]pyrene. Emission spectra are shown for BP in benzene (—), and as a suspension of crystals in aqueous buffer (---). After addition of DPPC vesicles to asbestos containing adsorbed BP (···) the emission spectra closely resembles that of BP in benzene.

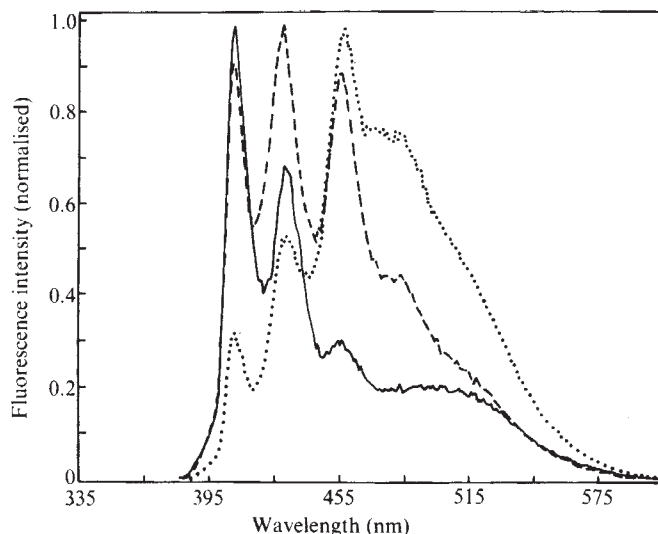
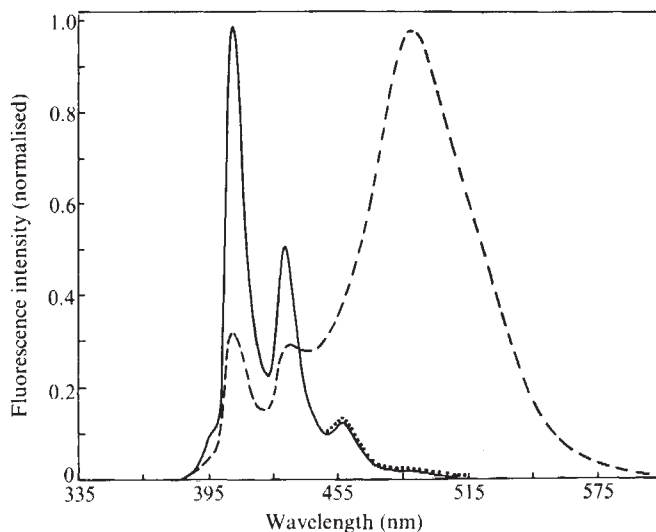


Fig. 2 Fluorescence emission spectra of benzo[a]pyrene. Normalised emission spectra for BP microcrystals (···), and for BP adsorbed to 8.6 mg silica (---) and 9.5 mg asbestos (—). In each case the sample was suspended in 10 ml of buffer, and contained 5 μ g BP.

Figure 2 shows the fluorescence emission spectra of BP when present as microcrystals which are suspended in aqueous buffer, and when the BP is adsorbed to silica or the amosite, which is the form of asbestos used in our studies (see Table 1 legend for experimental details). The emission spectra of the microcrystals and of the silica-adsorbed BP are similar, indicating the presence of BP microcrystals on the silica surface, or the presence of BP crystals which did not adsorb to silica during the solvent evaporation process. The emission spectra of the asbestos-adsorbed BP is dominated by the structured emission, indicating the presence of monomeric adsorbed BP. The superior ability of amosite to disperse BP is especially surprising when the surface areas of these two particulates are compared. As determined by nitrogen adsorption these surface areas are 5.7 and 381 $\text{m}^2 \text{g}^{-1}$ for asbestos and silica, respectively⁴⁻⁶.

The fluorescence emission spectra of BP are sensitive to both the weight percentage of carcinogen which is on the particulate and to the age of the samples. Over a period of weeks after preparation the emission spectra of BP on the silica samples becomes more structured and shows less excimer emission. We do not know what factors control this process, but temperature of sample storage and the presence of residual solvent do not seem to affect the rate of sample ageing. Less significant changes occur with asbestos-adsorbed BP. Irrespective of these considerations we consistently observed more BP excimer emission from the silica than from the asbestos samples. These observations include BP-to-particle weight ratios ranging from 3 to 0.1 mg BP per gram of particulate. For both particulates the excimer emission is less pronounced at the lower weight ratios (Table 1).

Emission spectra of aqueous suspensions of particles which contain BP retain a constant intensity and shape over a period of hours. On addition of phospholipid vesicles, an increase in fluorescence intensity occurs at 405 nm. Simultaneously the excimer emission disappears. After complete transfer of BP to the lipid bilayers the emission spectra of all samples listed in Table 1 become identical to that of BP in lipid bilayers (Fig. 1). Thus, significant decomposition of BP on the surface of the particles did not occur.

BP is only slightly soluble in water, so it seems likely that BP crystals will have a cellular availability which is limited by its rate of solubilisation in the aqueous phase. A particulate which adsorbs and disperses BP in a monomeric state could therefore facilitate BP solubilisation and hence increase cellular availability. The transfer of BP from the particulate's surface to the lipid

bilayer is conveniently quantitated by the increase in fluorescence intensity at 405 nm. The percentage of BP transferred is calculated from

$$\% \text{ BP transferred to membrane} = \frac{I(t) - I_0}{I_\infty - I_0}$$

where I_0 is the fluorescence intensities at 405 nm before addition of membranes, I_∞ is the intensity after complete transfer, and $I(t)$ the intensity at time t .

Figure 3 compares the membrane uptake rate of BP from the three states. As a model membrane we used single bilayer vesicles of dipalmitoyl-L- α -phosphatidylcholine^{7,8}. Asbestos-adsorbed BP shows the most rapid rate of membrane uptake, and the microcrystals of BP the slowest rate. Table 1 lists the time required for 50% transfer of BP for particulate samples with a range of BP-to-particulate weight ratios. Asbestos is more effective than silica at all weight ratios, with the most dramatic effect being observed at the lower ratios. Asbestos or silica, when added to preformed microcrystals of BP in buffer, did not increase the rate of BP membrane uptake or alter the shape of the fluorescence emission spectra of the suspended microcrystals. Therefore, only BP which is adsorbed to the particles displays an increased rate of membrane uptake. These rates increased as surface density of BP on the particulates decreased.

Our results clearly demonstrate that a particulate with a known co-carcinogenic effect is also effective in the dispersal and transport of benzo[a]pyrene. This observation may partially explain the co-carcinogenic effects of particulates and polynuclear aromatic hydrocarbons⁹⁻¹¹. Presumably a carcinogen must enter a cell before transformation, and such entry must involve transport of the polynuclear aromatic hydrocarbon across the plasma membrane. The lungs have mechanisms whereby particulates are cleared, these being phagocytosis by the lung macrophages and upward ciliary action followed by ingestion¹². Particulate matter which increases the rate of carcinogen transport could therefore increase the effective dose of carcinogen before clearance by these protective mechanisms. This speculation is supported by observations which indicate that soot particles which have been isolated from human lungs no longer contain BP¹³.

Particle-enhanced transport of carcinogens is unlikely to be the only mechanism of co-carcinogenesis. Studies with experimental animals have shown that the inhaled particulate haematite is co-carcinogenic with the injected systemic carcinogen diethylnitrosamine¹⁴. The tumours are localised in the lungs, which are the sites of particle deposition. However, particulate

matter is unlikely to facilitate carcinogen uptake in these conditions where exposure to particulates and carcinogens is by different routes of administration.

Epidemiological studies have correlated human exposure to asbestos with the incidence of mesothelioma, but this cancer

Table 1 Benzo[a]pyrene uptake by phospholipid vesicles from particulate adsorbed and microcrystalline states

mg BP per g particulate	Time for 50% uptake (min)		(Monomer emission) [†]	
	Silica	Asbestos	Silica	Asbestos
0.1	4.5	3	0.32	0.34
0.3	24	6.5	0.24	0.34
1.0	30	12	0.19	0.28
3.0	34	17	0.23	0.22

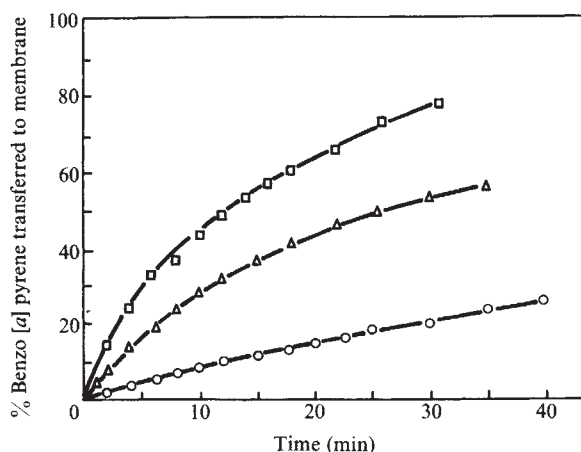
BP Microcrystals	Time for 15% uptake (min)		(Monomer emission)	
			(Total emission)	
No particles	23		0.17	
+ Amosite*	24		0.21	
+ Silica*	24		0.19	

Benzo[a]pyrene (BP) 99 + % purity was from Aldrich lot 101847 used without further purification. BP concentrations were obtained using a molar extinction coefficient of $57,500 \text{ M}^{-1} \text{ cm}^{-1}$ at 299.4 nm. Silica was obtained from preparative thin-layer chromatographic plates, $20 \times 20 \text{ cm}$, $1,000 \mu\text{m}$ thick (Analabs). The surface area of this sample, as determined by nitrogen adsorption, was $381 \text{ m}^2 \text{ g}^{-1}$. The particle size distribution is heterogeneous, with most particles being between 1.5 and $15 \mu\text{m}$ in diameter, and an average size of about $3 \mu\text{m}$ (ref 6). Amosite served as reference asbestos sample (International Union Against Cancer, Johannesburg). The surface area determined by nitrogen adsorption is $5.6 \text{ m}^2 \text{ g}^{-1}$. The particle size distribution is heterogeneous, ranging from 1.5 to $10 \mu\text{m}$, with the average size $2 \mu\text{m}$. For more detailed information on size and chemical composition see ref. 5. Particulate samples containing BP were prepared by mixing the particulate with a benzene solution of BP, followed by evaporation of the benzene under reduced pressure. 0.1 to 3.0 mg of BP was added for each g particulate. The volume of benzene added was always adequate to wet the entire particulate sample, typically 5 or 6 ml of benzene per g of particulate. The sample was rotated continuously during the benzene evaporation. After the sample appeared dry it was kept under vacuum for 30–45 min at 85°C . At the higher weight ratios BP crystals could be seen on the walls of the glass vessel. Any material which clung to the walls was discarded. After removal from the vessel the particulates were stored in the dark under an argon atmosphere. The amount of BP present on the particles was determined by solvent extraction using a Soxhlet extractor. Typically the extraction was continued for 4.5 h (about 50 individual extractions). All extractions were done in duplicate. The fluorescence emission spectra of the extracted material were identical to that of the starting material. Thus there seemed to be no degradation of BP on this particulate which interfered with our measurements. The amounts of BP adsorbed to silica and asbestos, were found to be $85 \pm 15\%$ of the amount added. Aqueous dispersions of BP crystals were prepared by evaporation of a benzene solution of BP to dryness, addition of buffer, followed by sonication for 30 min at 40 W using a Cole-Parmer Model 8845–2 water bath-type sonicator. We will refer to this preparation as being microcrystalline. In all studies the buffer was 0.01 M Tris-HCl, 0.05 M KCl, pH = 7.5. Phospholipid vesicles of dipalmitoyl-L- α -phosphatidylcholine were prepared by ultrasonic irradiation of the lipid suspended at 10 mg ml^{-1} . Sonication was performed using a Heat Systems Model 350 for 10 min at $\sim 50^\circ \text{C}$. Power input was $\sim 200 \text{ W}$. Vesicles were then centrifuged at $54,000 \text{ g}$ for 60 min. All fluorescence spectral data were obtained on an SLM Instruments, computerised photon counting spectrofluorometer (excitation wavelength, 296 nm; excitation and emission band-passes 8 and 4 nm, respectively; excitation filter, Corning 7–54; emission filters, 0–52 and 2 nm of 1 M NaNO₂). BP uptake was measured by the intensity increase at 405 nm. The emission filters do not transmit light below 395 nm. Without these filters we do observe a background signal from the asbestos and silica. Some background signal was seen above 395 nm, typically $<10\%$ of the total intensity. These backgrounds were quantitated using particles without BP, and subtracted from the spectra shown. Spectral data were obtained using a $2 \times 2 \text{ cm}$ cuvette, 4 cm high, which was thermostated and positioned on a magnetic stirrer so that the stirring bar stayed below the light path. Stirring was required to keep the particles or crystals in suspension. The exciting beam impinged upon the sample at an angle of 20° , and the emission observed from this illuminated surface was at 90° to the exciting light. For all the BP uptake kinetics reported here we used $5 \mu\text{g}$ of BP and 10 mg of DPPC. An amount of particulate containing $5 \mu\text{g}$ BP, or $5 \mu\text{g}$ BP crystals, was suspended in 10 ml of buffer. After measurement of the initial intensity 10 mg of DPPC were added in 1 ml of buffer to initiate the reaction. All measurements were performed at 25°C . Complete BP transfer to the bilayers was obtained by heating of the sample to 55°C for 30 min. The final fluorescence intensity was measured after equilibration at the experimental temperature of 25°C .

* Particles were added after formation of the microcrystals.

[†] This value was estimated from the ratio of area under the fluorescence emission spectrum from 380 to 440 nm to the total area from 380 to 605 nm. After complete uptake of BP by vesicles this ratio is 0.47.

Fig. 3 Rate of benzo[a]pyrene uptake in to lipid bilayers. The % of benzo[a]pyrene transferred to DPPC vesicles is shown for microcrystals of BP (○), silica adsorbed BP (△), and amosite adsorbed BP (□).



does not seem to be correlated with cigarette smoking¹⁵. Particulates such as fibrous glass¹⁶, asbestos^{10,11} and many other types of foreign bodies¹⁷ seem to be carcinogenic without deliberate exposure to chemical carcinogens. In fact, most types of cellular trauma and scarification will predispose the damaged areas to neoplastic transformation^{18,19}. Therefore it is clear that mechanisms other than particle enhanced transport are operative in carcinogenesis. The fluorescence methodology described here is useful with other polynuclear aromatic hydrocarbons⁶ and with nitrogen heterocyclic compounds²⁰. The ability to quantitate the delivery of carcinogens to cells should aid in elucidating the multiple etiology of carcinogenesis.

We thank the Freshwater Biological Research Foundation and especially its founder, Mr Richard Gray, Sr for assistance. This work was supported by the American Cancer Society (grant BC-261) and NIH (1R01-ES-1283). J.R.L. is an Established Investigator of the American Heart Association.

J. R. LAKOWICZ
J. L. HYLDEN

Freshwater Biological Institute
and Department of Biochemistry,
University of Minnesota, County Roads 15 & 19,
PO Box 100, Navarre, Minnesota 55392

Received 15 June; accepted 7 August 1978.

- Selikoff, I. J., Hammond, E. C. & Churg, J. *J. Am. med. Ass.* **204**, 104-110 (1968).
- Hammond, E. C. & Selikoff, I. J. *Int. Agency Res. Cancer Sci. Publ.* **8**, 312-317 (1973).
- Birks, J. B. & Cameron, J. W. *Proc. R. Soc. A* **249**, 297-317 (1959).
- Brunauer, S., Emmet, P. H. & Teller, E. *J. Am. chem. Soc.* **60**, 309-319 (1938).
- Timbrell, V. *Proc. Int. Pneumoconiosis Conf., Johannesburg, 1969*, 28-36 (Oxford University Press, 1970).
- Lakowicz, J. R., Englund, F. & Hidmark, A. *Biochim. biophys. Acta* (in the press).
- Huang, C. C. *Biochemistry* **8**, 344-352 (1969).
- Lawaczek, R., Kainosho, M. & Chan, S. I. *Biochim. biophys. Acta* **443**, 313-330 (1976).
- Saffioti, U., Cefis, F. & Kolb, L. H. *Cancer Res.* **28**, 104-124 (1968).
- Pylev, L. N. & Shabad, L. M. *Int. Agency Res. Cancer Sci. Publ.* **8**, 99-105 (1973).
- Pylev, L. N. *Bull. exp. Biol. Med. (USSR)* **52**, 1316-1319 (1961).
- Hatch, T. F. & Gross, P. *Pulmonary Deposition and Retention of Inhaled Aerosols* 69-85 (Academic, New York, 1964).
- Falk, H. L., Kotin, P. & Markal, I. *Cancer* **11**, 482-489 (1958).
- Nettesheim, P., Creasia, D. & Mitchell, T. T. *J. nat. Cancer Inst.* **55**, 688-691 (1975).
- Haley, T. J. *J. pharm. Sci.* **64**, 1435-1448 (1975).
- Stanton, M. F. & Wrench, C. J. *J. nat. Cancer Inst.* **45**, 797-821 (1972).
- Brand, K. G., Buoen, L. C. & Brand, I. J. *J. nat. Cancer Inst.* **55**, 319-322 (1975).
- Declman, H. T. *Proc. R. Soc. Med.* **20**, 19-20 (1927).
- Berenblum, I. *Arch. Path.* **38**, 233-244 (1944).
- Lakowicz, J. R., McNamara, M. & Steenson, L. *Science* **199**, 305-307 (1978).

Electromechanical properties of articular cartilage during compression and stress relaxation

THE main physiological function of articular cartilage is to act as a bearing material. It consists of an organic matrix made of collagen, proteoglycans, non-collagenous proteins, lipids and cells that are swollen with water^{1,2}. Its net fixed charge is negative and directly proportional in magnitude to its hexosamine content at each depth³. When cartilage is compressed, interstitial fluid is exuded. This flow of fluid with respect to the solid matrix plays an important part in controlling various rheological properties of the tissue, such as its creep deformation under load⁴⁻¹¹. Also, we have observed that measurable electrical potential differences are induced between the surface and deepest regions of cartilage when it is compressed. We report here our studies of this phenomenon to quantitatively relate the magnitude, sign and time dependence of the electrical potentials to the known features of cartilage mechanics and fluid flow¹¹, that is, in terms of tissue structure, function and composition. Our results suggest that an electrokinetic mechanism is primarily responsible for the mechanical-to-electrical transduction response.

Mechanically induced potentials in cartilage have been observed previously^{12,13} when a load was applied to one region of a hydrated specimen while electrodes were generally located

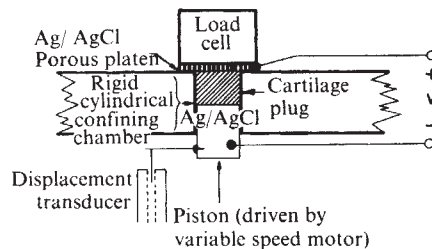


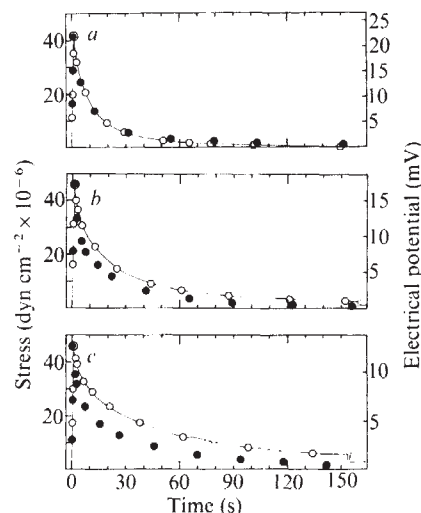
Fig. 1 Experimental configuration for measurement of electrical potential V_0 and mechanical stress σ due to an imposed strain ϵ and strain rate $\dot{\epsilon}$.

in an adjacent region. However, in these experiments the mechanical stress, strain and the fluid velocity field were so complex that it was difficult to relate quantitatively any induced electrical field to the mechanical fields. In our studies, therefore, a geometrical configuration was chosen so as to induce and measure electric fields that are colinear with the axial strain field. This enabled us to compare nonequilibrium electromechanical and mechanical rate (for example, transport) processes.

Circular disks of articular cartilage were compressed in a confining chamber with a chlorided silver rod against a porous Ag/AgCl platen (Fig. 1). The articular surface was positioned against the platen. During and after compression, the stress σ and the potential difference V_0 were measured. (The latter was monitored continuously with a Keithley 602 electrometer and a Tektronix 555 oscilloscope.) When final, time-invariant values of σ and V_0 were attained, the tissue was further compressed to a new strain, at the same strain rate $\dot{\epsilon} = 15\%$ per sec, and the procedure was repeated. Strains in the range of 5-20% were attained with each compression, corresponding to compressions of ~ 0.06 -0.2 mm. These strains resulted in peak stress levels for normal plugs that ranged over 250-650 p.s.i. (17 - 44×10^6 dyn cm⁻²). The apparatus was a modification of that used previously¹⁴ to study stress relaxation behaviour of articular cartilage.

To ensure that the measurements reflected the properties of the full thickness of the tissue, 40 plugs of cartilage and bone were cut from the parapatellar region of 14 medial femoral condyles of adult steers using a previously described procedure¹⁵. The underlying bone was then removed around the tide mark; the resulting cartilage plugs were 2.6 mm in diameter and ~ 1.0 -1.5 mm thick when fully hydrated. (However, plugs with a ≈ 1 mm region of bone gave similar results.) The plugs

Fig. 2 Potential (●) and stress (○) history for three successive compressions (a-c) of the same normal plug. ϵ , Cumulative strain: 12.7% (a), 26.4% (b) and 39.4% (c). $\dot{\epsilon}$, Strain rate = 15% s⁻¹ (a-c).



were prepared from tissue that was previously frozen. To assess the effect of such storage, five fresh plugs were prepared from one condyle and tested within 3 h of killing the animal, during which time the tissue was kept at 5 °C. Experiments were also carried out with nine plugs that had been partially digested with testicular hyaluronidase. The conditions used decreased the hexosamine content of the tissue by ~75% (ref. 16), and hence reduced its net negative charge. The plugs were equilibrated in 0.15 M NaCl at 4 °C for at least 1 h before use; longer equilibration times gave the same results. Also, one series of both the normal and enzymatically digested samples were equilibrated in 0.15 M NaCl at pH 2.8, and another series of normal plugs was equilibrated with 5 M NaCl at neutral pH.

The pertinent features of the electromechanical response are typified by the data of Figs 2 and 3. Figure 2 shows the time history of V_0 and σ for three successive compressions of a disk of cartilage equilibrated in 0.15 M NaCl at neutral pH. During the compression phase, $t < T$, $V_0(t)$ was positive and increased linearly with $\sigma(t)$ until time $t = T$ when compression was stopped. The ratio (V_0/σ) at $t = T$ always decreased monotonically with successive compressions at the same strain rate, $\dot{\epsilon}$ (Fig. 2). During the relaxation phase, $t > T$, both $V_0(t)$ and $\sigma(t)$ decreased; the decay with time of V_0 and σ was almost identical after the first compression. With each successive compression, these decay times were longer, but $\sigma(t)$ decayed at a successively slower rate than $V_0(t)$. Although the asymptotic stress was usually finite and progressively larger at the end of each successive compression relaxation cycle, V_0 typically relaxed to zero. These data were obtained with plugs that had been stored in the frozen state, but the results obtained with fresh samples were qualitatively identical.

The compression of enzymatically treated plugs equilibrated at 0.15 M NaCl, neutral pH, resulted in potentials positive in sign but reduced in magnitude by 45–75% (Fig. 3a). A similar result was obtained with normal plugs that had been equilibrated in 0.15 M NaCl, pH 2.8. However, all plugs that were enzymatically treated and equilibrated at pH 2.8 gave negative $V_0(t)$ values (Fig. 3b). Finally, the compression of normal plugs equilibrated in 5 M NaCl resulted in potentials that were drastically reduced in magnitude.

A comparison of the electrical and mechanical rate processes that were highlighted in this study suggests the importance of interstitial fluid flow in the observed electromechanical coupling phenomena. The movement of fluid and concomitant matrix reorientation is thought to control the stress relaxation of articular cartilage^{11,14}. Here, the sign and time dependence of V_0 on pH and neutral salt concentration suggest that the potentials are directly related to the flow of fluid and entrained counter-ions with respect to the net negatively charged solid matrix, that is, that an electrokinetic mechanism is dominant.

In the absence of other mechanisms, an electrokinetic model would predict the relaxation of the potential to zero on cessation of fluid flow, as was seen. Further, the observed decrease in (V_0/σ) at $t = T$ with increasing ϵ is consistent with the known¹⁷ increase in fluid flow resistance at higher ϵ . The instances of negative V_0 (Fig. 3b) would indicate that the combination of proteoglycan extraction and carboxyl group neutralisation changed the sign of the net matrix charge from negative to positive. Furthermore, the decrease in V_0 at high NaCl concentration is consistent with a screening of the effective fixed charge by mobile neutral salt ions, that is, a decrease of the electrokinetic ζ potential. Similar effects of pH and ionic strength have been seen with collagen membranes^{18,19}. The effect of pH on the dissociation of proteoglycan and collagen charged groups has also been reported^{20,21}. These results^{18–21}, together with the fact that streaming potentials have been induced *in vitro* using an applied pressure gradient but without varying the strain²², also support our conclusion.

Although electrochemical processes may contribute to the measured V_0 (for example, changes in interfacial Donnan equilibrium potentials, or nonequilibrium diffusion potentials due to strain-induced concentration gradients), their contribution was

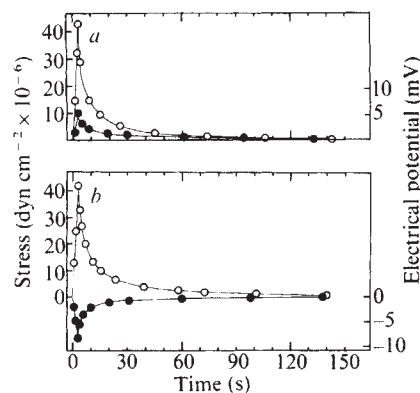


Fig. 3 Potential (●) and stress (○) history with *a*, enzymatically treated plug pre-equilibrated in 0.1 M NaCl at neutral pH ($\epsilon = 18\%$). *b*, Enzymatically treated plug pre-equilibrated in 0.1 M NaCl, pH 2.8 ($\epsilon = 34\%$). $\dot{\epsilon} = 15\% \text{ s}^{-1}$.

probably minimal. This emphasises the importance of considering the kinetics of $V_0(t)$ and $\sigma(t)$. For example, an estimate of the decay of a nonequilibrium concentration gradient and its concomitant diffusion potential, characterised by the diffusion time $\tau_{\text{diff}} \sim \delta^2/2D$, is almost two orders of magnitude slower than that observed for the decay of V_0 (based on a typical specimen thickness δ and internal diffusion coefficients D (ref. 23)). On the other hand, the Donnan interfacial potentials span a distance of a few Debye lengths²⁴. They can re-equilibrate in a time which is of the order of the Debye space charge relaxation time²⁵ (the interstitial fluid dielectric constant divided by its conductivity). This is a negligibly short time on the scale of $V_0(t)$, for relevant ranges of parameter values. On the other hand, the hypothesis that relative fluid flow is intimately linked not only to stress relaxation but to the induced potential V_0 is consistent with the observed similarity of the relaxation of $\sigma(t)$ and $V_0(t)$. This, in addition to our other results, lends further support to the notion that an electrokinetic effect is the dominant rate-limiting process.

The occurrence of mechanically induced electric fields *in vivo*, under equivalent physiological loads, may play an important part in regulating the transport of ions and interstitial fluid. We are now trying to assess quantitatively the influence of chemical modifications of the charge groups within the tissue in order to experimentally and theoretically characterise electromechanical influences on transport. Preliminary results show that compression of samples whose carboxyl groups were blocked (using established reactions²⁶) produces a V_0 that is consistently negative. Finally, we note that the successively differing relaxation rates found for stress as compared with potential (Fig. 2) may provide the first evidence and an important method for assessing the relative contributions of the fluid and solid phases to stress relaxation.

This work was begun while A.J.G. was visiting lecturer in the Department of Orthopaedic Surgery, Children's Hospital Medical Center and Harvard Medical School. The work was supported by Health Sciences Fund grant 77-02 (MIT), NSF grant ENG 77-06973 (MIT), NIH grant AM15671 (CHMC) and NSF grant 39017 (CHMC).

A. J. GRODZINSKY

Department of Electrical Engineering
and Computer Science,
Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139

H. LIPSHITZ
M. J. GLIMCHER

Department of Orthopaedic Surgery,
Children's Hospital Medical Center
and Harvard Medical School,
Boston, Massachusetts

Received 26 April; accepted 31 July 1978.

1. Serafini-Fracassini, A. & Smith, J. W. *The Structure and Biochemistry of Cartilage* (Churchill Livingstone, London, 1974).
2. Muir, H. in *Adult Articular Cartilage* (ed. Freeman, M. A. R.) Ch. 3 (Grune and Stratton, New York, 1973).
3. Maroudas, A. & Thomas, H. *Biochim. biophys. Acta* **215**, 216–219 (1970).
4. McCutchen, C. W. *Wear* **5**, 1 (1962).
5. Linn, F. C. & Sokoloff, L. *Arthritis Rheum.* **8**, 481 (1965).
6. Edwards, J. *Proc. Instn. mech. Engrs* **101**, 16 (1967).
7. Kempson, G. E., Freeman, M. A. R. & Swanson, S. A. V. *J. Biochem.* **4**, 239 (1971).
8. Hayes, W. C. & Mockros, L. F. *J. appl. Physiol.* **31**, 562 (1971).
9. Coletti, J. M., Akeson, W. H. & Woo, S. L. Y. *J. Bone Jt Surg.* **54A**, 147 (1972).
10. Torzilli, P. A. & Mow, V. C. *J. Biomech.* **9**, 541–552 (1975).
11. Mow, V. C. & Mansour, J. M. *J. Biomech.* **10**, 31 (1977).
12. Bassett, C. A. L. & Pawluk, R. J. *Science* **178**, 982 (1972).
13. Lotke, P. A., Black, J. & Richardson, S. *J. Bone Jt Surg.* **56A**, 1040 (1974).
14. Lipshitz, H., Mow, V. C., Torzilli, P. A., Eisenfeld, J. & Glimcher, M. J. *Proc. VII Int. Cong. Rheol.* 198 (1976).
15. Lipshitz, H. & Glimcher, M. J. *J. Biomech.* **7**, 293 (1974).
16. Lipshitz, H. & Glimcher, M. J. *Trans. orthop. Res. Soc.* **2**, 71 (1977).
17. Mansour, J. M. & Mow, V. C. *J. Bone Jt Surg.* **58A**, 509–516 (1976).
18. Grodzinsky, A. J. & Melcher, J. R. *IEEE Trans. Biomed. Eng.* **BME-23**, 421–433 (1976).
19. Grodzinsky, A. J. & Eisenberg, S. E. *J. Electrostatics* **5**, 35–53 (1978).
20. Freeman, W. D. S. C. & Maroudas, A. *Suppl. Ann. Rheum. Dis.* **34**, 44 (1975).
21. Bowes, J. H. & Kenten, R. H. *Biochem. J.* **8**, 575 (1968).
22. Maroudas, A. *Biophys. J.* **8**, 575 (1968).
23. Maroudas, A. *Biorheology* **12**, 233–248 (1975).
24. Plonsey, R. *Bioelectric Phenomena* (McGraw-Hill, New York, 1969).
25. Planck, M. *Ann. Physik. Chem.* **39**, 161 (1890).
26. Hoare, D. G. & Koshland, D. E. *J. biol. Chem.* **242**, 2447 (1967).

Naloxone reversal of endotoxin hypotension suggests role of endorphins in shock

It is known that the cardiovascular system is extremely sensitive to the effects of both exogenous^{1,2} and endogenous³ opiates. In rats, less than 1% of the morphine dose necessary to produce antinociception results in significant hypotension and bradycardia⁴. The endogenous opiate β -endorphin is stored along with pituitary adrenocorticotrophin (ACTH)^{5,6}, and the action of stressors seems to result in the release of both peptides^{5,6}. Therefore, it seemed likely that β -endorphin is released during shock states and that it might contribute to the hypotension. To test this hypothesis we used an endotoxin shock model^{7,8}. If endotoxin-induced hypotension were mediated through endorphin release, then blockade of endorphins should reverse such hypotension. Using the specific opiate antagonist, naloxone, we not only rapidly reversed endotoxin-induced hypotension, but also prophylactically blocked its occurrence. These findings suggest that endorphins may have a role in the pathophysiology of shock and that narcotic antagonists should be evaluated for their potential therapeutic value in the treatment of shock.

All studies were carried out using conscious rats with chronically indwelling arterial and venous catheters. Two days before experiments, male Wistar rats (Walter Reed strain) weighing 275–300 g were anaesthetised with pentobarbital 50 mg per kg. The tail artery was cannulated by the method of Chiueh and Kopin⁹ for subsequent measurement of blood pressure and heart rate. The external jugular vein was cannulated to allow for intravenous drug injections. Both cannulae were subcutaneously implanted so that they emerged at the back of the neck. The animals had access to food and water ad libitum and were maintained at 24 °C on a 12 h light–dark cycle.

Rats were studied in their home cages with arterial lines connected to a microtransducer (Narco Bio-Systems, model RP1500). Injections of naloxone hydrochloride (Endo Laboratories) or saline were made after administration of *Escherichia coli* lipopolysaccharide endotoxin (Difco, control no. 598159); all injections were given intravenously followed by a 0.3 ml saline intravenous flush to ensure that all of the drug was infused. Blood pressure and heart rate were continuously recorded on a Beckmann dynograph (type R).

Figures 1 and 2 represent the effects of treatment after injection of 4 mg endotoxin per rat. Usually within the first 5–15 min after endotoxin a precipitous fall in blood pressure occurred. When the mean arterial pressure (MAP) dropped to

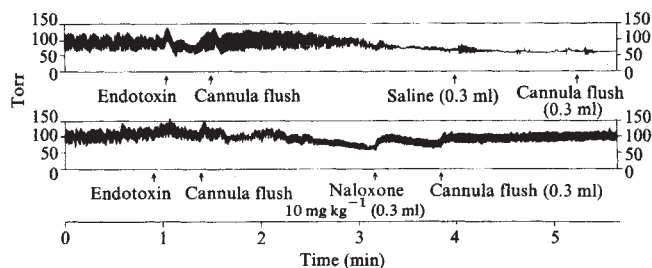


Fig. 1 Effects of intravenous saline (top) and naloxone (bottom) on the precipitous fall in blood pressure produced by 4 mg endotoxin in representative rats. Either saline (0.3 ml) or naloxone was injected when MAP fell to 65–70 mm Hg (torr). The prominent effect of the cannula flush (0.3 ml, bottom) is related to the clearing of residual naloxone from the line.

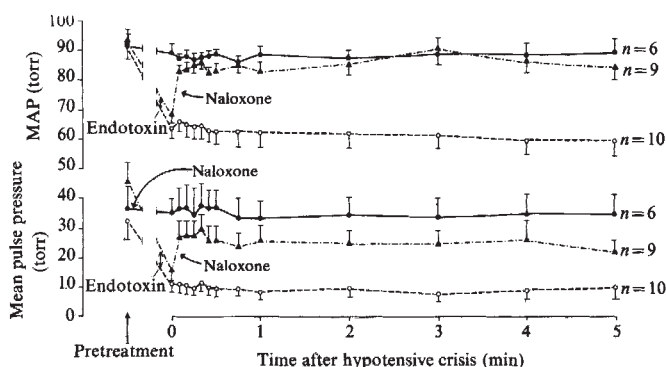
our pre-established criterion of 65–70 mm Hg (torr)⁸, an equivalent volume of either saline or naloxone 10 mg per kg was given by intravenous bolus.

Figure 1 shows the effects of saline and naloxone on endotoxin-induced hypotension in representative rats. In marked contrast to saline, which had only a minor transitory effect (Fig. 1, top), within 5 s after naloxone injection both MAP and pulse pressure were restored to pretreatment values (Fig. 1, bottom). The striking differences between naloxone-treated and saline control rats are further demonstrated by grouped data shown for the 5-min interval after endotoxin-induced hypotension (Fig. 2); again, a rapid return of MAP and pulse pressure towards control levels can be seen following naloxone injection. Naloxone in the absence of endotoxin (Fig. 2, solid line) was without effect on either MAP or pulse pressure.

A comparison of the effects of naloxone pretreatment with saline controls over a 2-h period after endotoxin administration is seen in Fig. 3. Due to the short biological half life of naloxone¹⁰, the initial 10 mg per kg dose of this drug was repeated every 30 min; saline controls were also treated at these same intervals. Although the same general pattern of changes was observed in both groups (Fig. 3), the saline control animals exhibited two hypotensive episodes (10 and 45 min after endotoxin administration) during which blood pressure values were significantly lower than before endotoxin (repeated measurement analysis of variance using Hotelling's T^2 statistic¹¹, $F = 12.29$, $P = 0.003$). By contrast, naloxone-pretreated animals did not show significant declines in blood pressure after endotoxin administration ($F = 0.83$, $P = \text{NS}$).

The reversibility of endotoxin-induced hypotension over a 2-h period is represented in Fig. 4. After injection of 8 mg endotoxin, animals were randomly divided into two equal

Fig. 2 Results of various drug injections on mean arterial blood pressure (top) and mean pulse pressure (bottom). Measurements at $t = 0$ began 5 s after intravenous catheter flush. Naloxone alone (●) did not significantly change either measure. Endotoxin 4 mg intravenously resulted in a significant drop in MAP and pulse pressure from control levels (Student's t test, $P < 0.01$ both groups). Subsequent naloxone injection reversed the drop in MAP and pulse pressure (▲), whereas saline injections (○) were without effect. Points represent averages $\pm$ s.e.m. Naloxone doses 10 mg per kg.



groups. One-half received naloxone 10 mg per kg and the other half saline; these were given only after individual MAPs had dropped to 65–70 torr. Animals receiving naloxone after endotoxin-induced hypotension had significantly higher MAP than those receiving saline (analysis of variance¹², $F = 10.55$, $P = 0.009$).

In addition to blood pressure effects, heart rate was also measured. In the saline control group, heart rate data tended to parallel the blood pressure changes, the lowest average heart rate being observed 55 min after injection (347 ± 33 beats per min, $n = 4$) followed by an increase to significantly higher values at 120 min (453 ± 17 beats per min, $n = 4$, Student's $t = 2.86$, $P < 0.02$). In naloxone-pretreated rats, subsequent injection of endotoxin had no effect on heart rate.

The results of this study clearly demonstrate that naloxone not only blocks but also reverses the hypotensive effects of endotoxin. Heart rate changes after endotoxin also seem to be

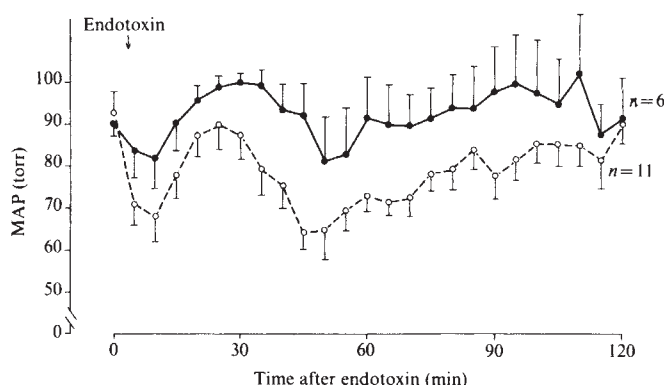


Fig. 3 Effects of naloxone pretreatment on 2 h MAP measurements after endotoxin. Pretreatment with naloxone 10 mg per kg (●) blocked the effects of 4 mg endotoxin on MAP. The groups treated with 4 mg endotoxin alone (○) had significantly lower MAP values compared with pre-endotoxin control values (see text). Vertical bars are $\pm$ s.e.m.

blocked by naloxone. During these experiments the administration of naloxone was discontinued after 2 h (Fig. 3), and survival was not affected. Studies are in progress to evaluate the effects of more prolonged naloxone administration.

The efficacy of the specific opiate antagonist, naloxone, as a modifier of physiological and behavioural events has been used as a tool to infer endorphin involvement in various systems^{13–15}. From that perspective, the effects of naloxone in these present studies implicate endorphin as a factor in endotoxin shock. Interestingly, it has often been suggested that bacterial endotoxins result in the secondary release of endogenous substances which, in turn, may cause some of the cardiovascular effects associated with this form of shock^{7,8}. Our studies suggest that endorphin may be one of these postulated substances.

Endotoxins are known to result in the secretion of ACTH from the pituitary, presumably by activation of hypothalamic releasing factors¹⁶. It has been shown recently that both ACTH and β -endorphin are stored in common pituitary sites and concomitantly regulated^{5,6}. It seems probable that endotoxins would therefore also release pituitary β -endorphin. Moreover, it seems likely that the established therapeutic efficacy of synthetic corticosteroids^{8,17} in the treatment of shock may be a partial consequence of their inhibitory effects on pituitary β -endorphin release⁵. The enhanced sensitivity of adrenalectomised animals to shock states¹⁸ as well as to intravenous β -endorphin injections¹⁹, both of which are attenuated by corticosteroid pretreatment, are also consistent with the hypothesis that endorphin release is important in the pathophysiology of shock.

The effect of naloxone in reversing the early hypotension of endotoxin shock suggests that naloxone may be efficacious in the

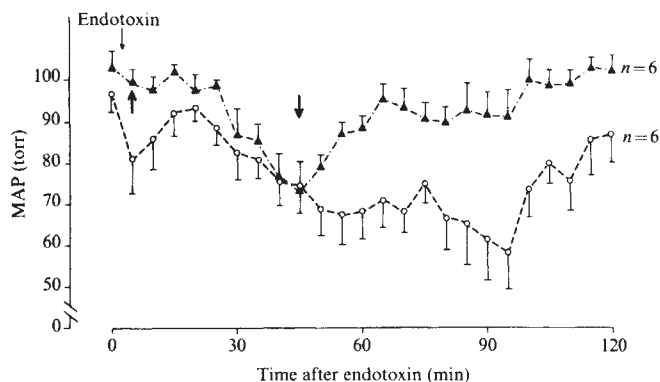


Fig. 4 Reversal of endotoxin shock by naloxone. After injection of 8 mg endotoxin, rats were injected with naloxone 10 mg per kg (▲) or saline (○) after MAP reached 65–70 torr. Large arrows indicate approximate times of naloxone injection. Some animals required two naloxone injections because the hypotensive effects produced by endotoxin were biphasic. Naloxone injections significantly improved MAP (see text). Vertical bars are $\pm$ s.e.m.

treatment of septic shock. The possibility that the endogenous opiates may also be involved in the pathophysiology of other shock models is currently being investigated in our laboratory.

We thank Drs M. Woods and J. L. Meyerhoff for advice and Dr T. Jacobs, Dr R. McCarty and J. Quinn for their help.

J. W. HOLADAY
A. I. FADEN

Department of Medical Neurosciences,
Division of Neuropsychiatry,
Walter Reed Army Institute of Research,
Washington DC 20012

Received 15 June; accepted 31 July 1978.

1. Evans, G. J., Nasmyth, P. A. & Stewart, H. C. *Br. J. Pharmac.* **7**, 542–552 (1952).
2. Gokhale, S. D., Gulati, O. D. & Joshi, N. Y. *Br. J. Pharmac.* **21**, 273–284 (1963).
3. Florey, J. & Mediavilla, A. *Brain Res.* **138**, 585–590 (1977).
4. Fennessy, M. R. & Rattray, J. F. *Eur. J. Pharmac.* **14**, 1–8 (1971).
5. Guillemin, R. *et al. Science* **197**, 1367–1369 (1977).
6. Rossier, J., French, E. D., Rivier, C., Ling, N., Guillemin, R. & Bloom, F. E. *Nature* **270**, 618–620 (1978).
7. Waisbren, B. A. *Am. J. Med.* **36**, 819–824 (1964).
8. Gilbert, R. P. *Physiol. Rev.* **40**, 245–279 (1960).
9. Chiu, C. C. & Kopin, I. J. *Am. J. Physiol.* **234**, H-690–695 (1978).
10. Berkowitz, B. A., Ngai, S. H., Hempstead, J. & Spector, S. J. *Pharmac. exp. Ther.* **195**, 499–504 (1975).
11. Morrison, D. F. in *Multivariate Statistical Methods* 133–141 (McGraw-Hill, New York, 1967).
12. Winer, B. J. in *Statistical Principles in Experimental Design* 514–539 (McGraw-Hill, New York, 1971).
13. Akil, H., Madden, J., Patrick, R. L. & Barchas, J. D. in *Opiates and Endogenous Opioid Peptides* (ed. Kosterlitz, H. W.) (North-Holland, New York, 1976).
14. Walker, J. M. *et al. Science* **196**, 85–87 (1977).
15. Holaday, J. W., Wei, E., Loh, H. H. & Li, C. H. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2923–2927 (1978).
16. Moberg, G. P. *Am. J. Physiol.* **220**, 397–400 (1971).
17. Nies, A. S. in *Clinical Pharmacology* (eds Melmon, K. L. & Morelli, H. F.) (Macmillan, New York, 1972).
18. Zweifach, B. W. & Thomas, L. J. *exp. Med.* **406**, 385–401 (1957).
19. Holaday, J. W., Law, P.-Y., Tseng, L.-F., Loh, H. H. & Li, C. H. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4628–4632 (1977).

Nicotinic cholinergic stimulation increases cyclic GMP levels in vertebrate skeletal muscle

CYCLIC GMP seems to be an intracellular second messenger for certain actions of acetylcholine at muscarinic receptors (for a review, see ref. 1). For example, previous studies have provided evidence that activation of muscarinic, but not of nicotinic, cholinergic receptors leads to an increase in the level of cyclic GMP in many tissues^{2–8}, including smooth muscle^{3–5}. Also, it has been demonstrated that contraction of cross-striated muscle fibres of the giant barnacle elicited by nerve stimulation is associated with an increase in the level of muscle cyclic GMP⁹.

These observations prompted us to determine whether a similar cyclic GMP increase is associated with contractile activity in vertebrate skeletal muscle. We report here that activation of a nicotinic cholinergic synapse raises the level of cyclic GMP in frog skeletal muscle. As in smooth^{4,5} and in barnacle⁹ muscle, this cyclic GMP increase in vertebrate skeletal muscle is a consequence of a neurotransmitter-induced depolarisation of the muscle membrane.

The gastrocnemius muscles of 3–4 inch *Rana pipiens* (northern variety) were used in this study. The cyclic nucleotide content of resting muscles from different frogs varied several-fold, but the levels in the two muscles from a single frog were essentially equal (Table 1, line 1). Therefore, in the experiments described below, one gastrocnemius from each frog served as the test muscle and the other as the control.

When the motor nerve supplying a test muscle was stimulated for 5, 15 or 30 s, in the conditions described in the legend to Table 1, both a maintained tetanic contracture and an increase in the level of muscle cyclic GMP resulted. The increase seemed to be maximal after 15 s, at which time test muscle cyclic GMP concentration was 230% that of the paired control muscle. The effect of stimulation for periods longer than 30 s was not determined because the test muscle fatigued rapidly beyond this point. In agreement with other reports¹⁰, nerve stimulation did not significantly increase the level of cyclic AMP in frog skeletal muscle.

Table 1 Effect of nerve stimulation on cyclic nucleotide levels in frog skeletal muscle

Duration of stimulation (s)	Cyclic GMP % control	Cyclic AMP
Unstimulated	101 ± 7 (9)	105 ± 8 (13)
5	162 ± 21 (7)*	129 ± 20 (7)
15	230 ± 45 (19)†	117 ± 14 (14)
30	192 ± 23 (7)‡	117 ± 13 (6)

Paired gastrocnemii, with nerve supply intact, were removed from each frog and preincubated for 1 h at room temperature in frog saline of the following composition: 111 mM Na⁺, 3 mM K⁺, 1.8 mM Ca²⁺, 117.6 mM Cl⁻, 5 mM MOPS (3-(*N*-morpholino)propanesulphonic acid), adjusted to pH 7.4 with NaOH. One gastrocnemius from each frog was arbitrarily chosen as the test muscle and the other as the control muscle. After the preincubation period, the motor nerve supplying the test muscle was stimulated by a suction electrode with supramaximal pulses of 0.2 ms duration delivered at 30 Hz for 5, 15 or 30 s. Following this stimulation, the control and test muscles were rapidly homogenised in 4.0–7.0 ml of ice-cold 7% trichloroacetic acid with an Ultraturrax. The homogenates were then centrifuged at 27,000 *g* for 5 min. The protein content of the pellets was determined by the method of Lowry *et al.*¹⁸. 3.0–3.5-ml aliquots of the supernatants were extracted five times with 10 ml of ether, lyophilised and reconstituted in 0.5 ml of 50 mM sodium acetate–2 mM EDTA (pH 6.2). The cyclic GMP content of the reconstituted samples was measured by the radioimmunoassay method of Steiner *et al.*¹⁹, and the cyclic AMP content by the method of Brown *et al.*²⁰. The recovery of exogenous cyclic GMP and cyclic AMP, added to whole muscle homogenates, was 83% and 89%, respectively. When sodium acetate-reconstituted samples were treated with beef heart phosphodiesterase (Boehringer), the cyclic nucleotide content was reduced to undetectable levels. The cyclic GMP and cyclic AMP concentrations of resting muscles were 0.056 ± 0.006 (mean ± s.e.m.; 28 muscles) and 1.13 ± 0.07 (mean ± s.e.m.; 26 muscles) pmol per mg protein, respectively. The protein levels of the two muscles from a given frog were essentially equal; test muscle protein was 102 ± 2% (mean ± s.e.m.; 22 muscle pairs) that of control muscle. For each muscle pair, results were calculated as (100 (cyclic nucleotide content of test muscle in pmol per muscle)/(cyclic nucleotide content of control muscle in pmol per muscle)). The values represent the mean % ± s.e.m. of these data; the numbers in parentheses indicate the number of muscle pairs used in each experiment. Statistical significance was calculated using a two-tailed Student's *t*-test.

* Significantly different from unstimulated, *P* < 0.05.

† *P* < 0.02.

‡ *P* < 0.01.

Table 2 shows the effect of nicotinic and muscarinic cholinergic antagonists on the response of cyclic nucleotides to nerve stimulation. Muscle pairs were preincubated for 1 h in frog saline containing the antagonists, and the motor nerve of the test muscle then stimulated for 15 s. Both nicotinic antagonists, curare (5 × 10⁻⁴ M) and α-bungarotoxin (10⁻⁶ M), blocked the nerve-evoked contraction almost completely, and both compounds abolished the elevation of cyclic GMP in response to nerve stimulation (Table 2, lines 3, 4). The muscarinic antagonist scopolamine (5 × 10⁻⁴ M) did not visibly affect the nerve-evoked contraction; nerve stimulation in the presence of scopolamine increased the cyclic GMP level of the test muscle to 190% that of the paired control muscle (Table 2, line 5). The difference between the nerve-stimulated increase in cyclic GMP in the presence and in the absence of scopolamine was not statistically significant (compare lines 2 and 5, Table 2; *P* > 0.1 by Student's *t*-test). These data indicate that the nerve-stimulated increase in cyclic GMP resulted from the activation, by endogenous acetylcholine, of nicotinic, rather than muscarinic, receptors.

To determine whether, as in barnacle and smooth muscle, membrane depolarisation, without neurotransmitter receptor activation, is able to increase cyclic GMP in frog skeletal muscle, muscle pairs were preincubated in frog saline for 1 h after which the test muscle was caused to contract by direct electrical stimulation for 15 s (as described in the legend to Table 2). The isometric tension response of the muscle to direct electrical stimulation was similar to that observed during stimulation of the motor nerve. The cyclic GMP content of the directly stimulated test muscle was 227% that of the control muscle (Table 2, line 6). There was also a slight increase in cyclic AMP in these directly stimulated muscles, but this increase was not statistically significant. The cyclic GMP increase elicited in response to direct electrical stimulation was not due to activation of nicotinic receptors by endogenous acetylcholine released from motor nerve terminals, as this increase was not blocked by the nicotinic antagonist α-bungarotoxin. Thus, a 1-h preincubation in 10⁻⁶ M α-bungarotoxin, which inhibited neuromuscular transmission, did not visibly affect the electrically stimulated contracture, and direct stimulation in the presence of α-bungarotoxin elevated the level of test muscle cyclic GMP to 192% that of the paired control muscle (Table 2, line 7; lines 6 and 7 not statistically different by Student's *t*-test, *P* > 0.1). These data suggest that depolarisation of the muscle membrane in response to activation of acetylcholine receptors, is a step in the sequence of events by which neuromuscular activity elevates the cyclic GMP level of frog skeletal muscle.

Previous studies had demonstrated that activation of muscarinic, but not nicotinic, cholinergic receptors causes an increase in the level of cyclic GMP in many target tissues; the tissues studied included heart^{2,3}, smooth muscle^{3–5}, brain³, salivary gland⁴, sympathetic ganglia⁶ and exocrine pancreas^{7,8}. The present study of frog skeletal muscle represents the first demonstration of a nicotinic cholinergic-mediated increase in cyclic GMP. Previous failures to observe nicotinic-stimulated increases in cyclic GMP may be explained partly by the fact that many of the tissues examined lack nicotinic receptors and partly by the fact that bath-applied agonists rapidly desensitise the nicotinic, as opposed to the muscarinic, receptor (for a review, see ref. 11).

The present results on vertebrate skeletal muscle and previous studies on invertebrate cross-striated muscle⁹ and vertebrate smooth muscle^{4,5} indicate that, in all three types of muscle, contractile activity is associated with an increase in cyclic GMP. In each case, contractile stimuli which depolarise the muscle membrane, without activating neurotransmitter receptors, also increase muscle cyclic GMP levels; these results suggest that the neurotransmitter increases cyclic GMP levels at least in part indirectly through membrane depolarisation, rather than through direct activation of a neurotransmitter-sensitive guanylate cyclase. Other evidence, obtained in barnacle muscle⁹ and in smooth muscle^{4,5}, suggests that the mechanism by which

Table 2 Effect of cholinergic antagonists on cyclic nucleotide levels in stimulated frog skeletal muscle

Preincubation solution	Mode of stimulation of test muscle	Cyclic GMP % control	Cyclic AMP % control
1. Normal saline	None	101 ± 7 (9)	105 ± 8 (13)
2. Normal saline	Nerve stimulation	230 ± 45 (19)*	117 ± 14 (14)
3. 5×10^{-4} M curare-saline	Nerve stimulation	104 ± 13 (11)	97 ± 10 (6)
4. 10^{-6} M α -bungarotoxin-saline	Nerve stimulation	101 ± 16 (5)	112 ± 14 (5)
5. 5×10^{-4} M scopolamine-saline	Nerve stimulation	190 ± 14 (8)†	118 ± 12 (8)
6. Normal saline	Direct stimulation	227 ± 22 (22)†	138 ± 17 (10)
7. 10^{-6} M α -bungarotoxin-saline	Direct stimulation	192 ± 29 (10)*	111 ± 11 (10)

Paired gastrocnemii, with nerve supply intact, were removed and preincubated for 1 h at room temperature in the indicated solutions. 'Nerve stimulation' indicates that the nerves supplying the test muscles were stimulated for 15 s as described in the legend to Table 1. 'Direct stimulation' indicates that test muscles were placed in a plexiglass chamber and a baffle was fitted around the muscles to divide the chamber into two compartments; supramaximal pulses (1-ms duration; 30 Hz) were applied for 15 s between the two compartments. The values represent the cyclic nucleotide content of the test muscle as the mean % $\pm$ s.e.m. of that of the paired control muscle. The numbers in parentheses indicate the number of muscle pairs used in each experiment. Statistical significance was calculated using a two-tailed Student's *t*-test.

* Significantly different from unstimulated (line 1), $P < 0.02$.

† $P < 0.01$.

depolarisation increases cyclic GMP involves an increase in the intracellular concentration of free calcium. A similar mechanism may be responsible for the cyclic GMP increase in frog skeletal muscle.

The role of cyclic GMP in muscle function remains unknown. Ong and Steiner¹², using an immunocytochemical procedure, showed that cyclic GMP is predominantly located over the A bands in unstimulated mammalian skeletal muscle, but the functional significance of this observation has not yet been determined. Cyclic GMP may be part of a positive or negative feedback system for regulating the concentration of free calcium in muscle^{13,14}. Another possibility is that cyclic GMP mediates some of the effects of contractile activity on muscle physiology. For example, contractile activity increases the glucose permeability of vertebrate skeletal muscle¹⁵, and can both maintain muscle mass¹⁶ and prevent the appearance of extra-junctional acetylcholine receptors¹⁷ in denervated mammalian muscle. It is possible that cyclic GMP is the intracellular mediator of one or more of these effects. The observation that contractile stimuli elevate cyclic GMP in several types of muscle indicates that this cyclic GMP increase is of physiological importance.

We thank Dr Thomas Lentz for purified α -bungarotoxin and Dr Robert Study for participating in the early experiments of this investigation and for providing the materials for the cyclic GMP radioimmunoassay. This research was supported by the USPHS.

ERIC J. NESTLER
KURT G. BEAM*
PAUL GREENGARD

Department of Pharmacology,
Yale University School of Medicine,
New Haven, Connecticut 06510

Received 26 June; accepted 31 July 1978.

* Present address: Department of Physiology and Biophysics, University of Iowa School of Medicine, Iowa City, Iowa 52242.

- Goldberg, N. D., O'Dea, R. F. & Haddox, M. D. *A. Rev. Biochem.* **46**, 823–896 (1977).
- George, W. J., Polson, J. B., O'Toole, A. G. & Goldberg, N. D. *Proc. natn. Acad. Sci. U.S.A.* **66**, 398–403 (1970).
- Lee, T. P., Kuo, J. F. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3287–3291 (1972).
- Schultz, G., Hardman, J. G., Schultz, K., Baird, C. E. & Sutherland, E. W. *Proc. natn. Acad. Sci. U.S.A.* **70**, 3889–3893 (1973).
- Clyman, R. I., Sandler, J. A., Manganiello, V. C. & Vaughan, M. *J. clin. Invest.* **55**, 1020–1025 (1975).
- Kebabian, J. W., Steiner, A. L. & Greengard, P. *J. Pharmac. exp. Ther.* **193**, 474–488 (1975).
- Haymovits, A. & Scheele, G. *Proc. natn. Acad. Sci. U.S.A.* **73**, 156–160 (1976).
- Christophe, J. P., Frandsen, E. K., Conlen, T. R., Krishna, G. & Gardner, J. D. *J. biol. Chem.* **251**, 4640–4645 (1976).
- Beam, K. G., Nestler, E. J. & Greengard, P. *Nature* **267**, 534–536 (1977).
- Posner, J. B., Stern, R. & Krebs, E. G. *J. biol. Chem.* **240**, 982–985 (1965).
- Steinbach, J. H. & Stevens, C. F. in *Frog Neurobiology* (eds Llinas, R. & Precht, W.) 33–92 (Springer, Berlin, 1976).
- Ong, S. H. & Steiner, A. L. *Science* **195**, 183–185 (1977).
- Nilsson, K. B. & Andersson, R. G. *Acta physiol. scand.* **99**, 246–253 (1977).
- Schultz, K. D., Schultz, K. & Schultz, G. *Nature* **265**, 750–751 (1977).
- Holloszy, J. O. & Narahara, H. T. *J. biol. Chem.* **240**, 3493–3500 (1965).
- Westgaard, R. H. *J. Physiol., Lond.* **251**, 683–697 (1975).
- Lomo, T. & Rosenthal, J. *J. Physiol., Lond.* **221**, 493–513 (1972).
- Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265–275 (1951).
- Steiner, A. L., Parker, C. W. & Kipnis, D. M. *J. biol. Chem.* **247**, 1106–1113 (1972).
- Brown, B. L., Albano, J. D. M., Ekins, R. P., Sgherzi, A. M. & Tampion, W. *Biochem. J.* **121**, 561–562 (1971).

Effect of guanine nucleotides on striatal dopamine receptors

GUANINE NUCLEOTIDES have been shown to regulate the sensitivity of adenylate cyclase to hormones in several systems^{1–5}, including the dopamine-sensitive adenylate cyclase in the caudate nucleus^{2,4}. It has also been reported that in the presence of GTP the affinities of glucagon receptors⁶, β -adrenergic receptors^{7,8}, prostaglandin E_1 receptors⁹ and opiate receptors¹⁰ for agonists are decreased. For β -adrenergic receptors this effect of GTP has been described as 'agonist-specific', as it is seen with agonists but not with antagonists^{7,8}. We report here a similar effect of GTP on the ability of dopamine-receptor agonists to compete for ³H-spiroperidol binding sites on rat striatal membranes. The presence of 0.3 mM GTP led to a four-to-fivefold increase in the K_d values for the inhibition of ³H-spiroperidol binding by dopamine-receptor agonists. No changes in K_d values were observed for antagonists. Dopamine-receptor agonists and antagonists have been defined by the stimulation or inhibition of dopamine-sensitive adenylate cyclase in the neostriatum^{11,12} or by their effects on the canine renal artery¹³. Controversy exists, however, as to whether binding studies using labelled neuroleptics such as spiroperidol actually measure dopamine receptors. Our finding of an agonist-specific effect of GTP supports the conclusion that ³H-spiroperidol is binding to functional dopamine receptors in the caudate nucleus.

The binding of the potent dopamine receptor antagonist ³H-spiroperidol was studied *in vitro* using a direct filtration binding assay (see legend to Fig. 1a and refs 14–16). At 37°C ³H-spiroperidol bound to rat striatal membranes with an observed affinity of approximately 80 pM, reached equilibrium levels in 8 min, and had a half-time of dissociation of 20 min. Specific binding was defined as the difference between ³H-spiroperidol bound in the absence and presence of 2 μ M (+)butaclamol; this constituted 90–95% of total binding. At the same concentration, (–)butaclamol was inactive. GTP had no effect on the binding characteristics of ³H-spiroperidol. Scatchard analysis was consistent with the existence of only a single

class of high-affinity binding site. The presence of GTP did not lead to changes in either the K_d (78 ± 6 pM and 85 ± 5 pM in the absence, and presence of GTP, respectively; $n = 6$) or the density of binding sites for ^3H -spiroperidol (305 ± 1 fmol per mg and 306 ± 1 fmol per mg in the absence and presence of GTP, respectively). The control K_d value in this study was approximately fourfold lower than that reported previously^{14,15}. This is probably due to species differences—Creese *et al.* used bovine caudate¹⁴—and to the use of different tissues—Fields *et al.* used whole rat brain¹⁵.

The addition of GTP led to a decrease in the apparent affinity of dopamine for the ^3H -spiroperidol binding site ($\text{EC}_{50} = 0.003$ mM GTP). The effect of GTP on rat striatal membranes was maximal at the concentration used (0.3 mM GTP). At this high concentration the order of efficacy of purine nucleotides was as follows: $\text{GTP} = \text{GDP} > \text{ITP} \gg \text{GMP} = \text{guanosine} = \text{ATP} = \text{control}$ (Fig. 1a). These results are similar to the effects of purine nucleotides on β -adrenergic receptors⁷.

A variety of dopamine receptor agonists and antagonists effectively competed with ^3H -spiroperidol for binding sites on

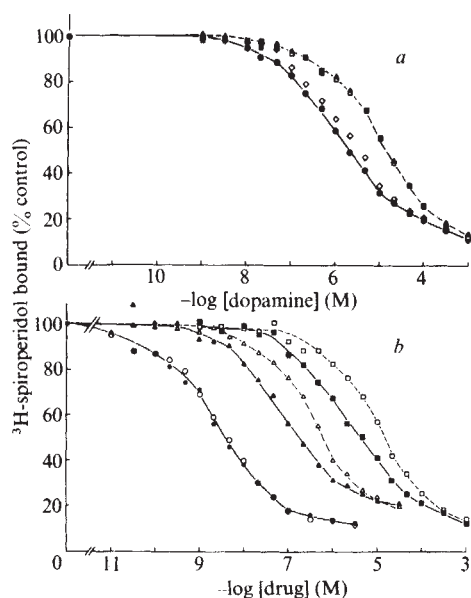


Fig. 1 a, Effect of guanine nucleotides on the inhibition of ^3H -spiroperidol binding by dopamine in rat striatal membranes. Striatal membranes were isolated following homogenisation in 200 volumes of 0.32 M sucrose/10 mM Tris buffer (pH 7.5) by centrifugation at 20,000 g for 10 min at 4°C. Radioligand, drugs and nucleotides were dissolved in 2.8 mM ascorbic acid containing 10 μg per ml bovine serum albumin (BSA). An aliquot of the resuspended membranes (0.1 mg protein in 850 μl of 0.9% NaCl/10 mM Tris buffer, pH 7.5) were added to tubes containing 50 μl ^3H -spiroperidol (26 Ci mmol^{-1} ; 0.4 nM final concentration), 50 μl drug and 50 μl ascorbate/BSA or nucleotide (0.3 mM final concentration). Samples were incubated at 37°C for 10 min; the reaction was then terminated by the addition of 10 ml ice-cold buffered saline and samples were immediately filtered through type-AE glass fibre filters. Each filter was washed at room temperature with an additional 10 ml of buffered saline, and radioactivity was determined by liquid scintillation spectrometry in 7 ml of a Triton X-100/toluene-based fluor. The counting efficiency for tritium was approximately 32%. Points on the displacement curves represent the mean values of three independent determinations. In these experiments ^3H -spiroperidol binding ranged from 800 c.p.m., in the absence of dopamine, to 80 c.p.m., in the presence of 1 mM dopamine. Standard errors (not shown) ranged over 5–15%. ●, Control; ▲, GTP; □, GDP; ◇, GMP. b, The effect of 0.3 mM GTP on the inhibition of ^3H -spiroperidol binding by dopamine-receptor agonists and antagonists in rat striatal membranes. Methods used were identical to those described above. ●, α -flupenthixol; ○, α -flupenthixol + GTP; ▲, ($\pm$)ADTN; △, ($\pm$)ADTN + GTP; ■, epinine; □, epinine + GTP.

striatal membranes (Fig. 1b and Table 1). Addition of 0.3 mM GTP to the incubation medium did not affect the inhibition of ^3H -spiroperidol binding by the potent antagonist α -flupenthixol. In the presence of GTP, however, the agonists ($\pm$)amino-6,7-dihydroxy-1,2,3,4-tetrahydronaphthalene (ADTN) and epinine were less potent in competing for ^3H -spiroperidol binding sites (Fig. 1b). Several dopamine-receptor agonists and the partial agonist apomorphine were less potent in displacing ^3H -spiroperidol binding in the presence of GTP, whereas antagonists were not affected by the presence of GTP (Table 1). The magnitude of the effect was independent of the potency of the agonist tested. Hill coefficients for the agonists/partial agonist were less than unity (0.49–0.61; Table 1) and increased only slightly on the addition of GTP (0.59–0.76). This finding differs from that with β -adrenergic receptors, in which GTP leads to a marked increase in the Hill coefficients for agonists towards unity. In the case of β -adrenergic receptors, Hill coefficients for antagonists are approximately 1.0 in either the presence or absence of GTP⁵.

It should be pointed out that the active isomers of the dopamine-receptor antagonists have Hill coefficients of approximately 0.65, whereas the inactive isomers have Hill coefficients of about 1.0 (Table 1). This difference suggests that the effect of the less active isomers is a reflection of their intrinsic activity and not of contamination with a small amount of the active isomers. It is likely that there are differences in the site or sites occupied by the active and inactive isomers. The Hill coefficients of the antagonists were not changed by the addition of GTP. Low Hill coefficients (<1) could indicate negative cooperativity between individual receptor sites or the existence of multiple binding sites, although it is not known whether either of these explanations is correct.

A possible explanation for the agonist-specific effect of GTP comes from studies of the binding of labelled agonists to glucagon receptors^{1,6}, β -adrenergic receptors⁸, prostaglandin E_1 receptors⁹ and opiate receptors¹⁰. In these systems, guanine nucleotides increase the rate of dissociation of agonists from their binding sites. Agonists, therefore, seem to be less potent in the presence of GTP. Labelled agonists for dopamine receptors have been used in direct binding studies^{17–19}. However, the specific binding of ^3H -dopamine to rat striatal membranes is usually only about 20% of total binding¹⁸. We have thus far been unable to examine the effects of GTP on ^3H -dopamine binding in the rat striatum.

Several reports have suggested that the action of GTP may be involved in the coupling of the receptor with adenylate cyclase^{1,5,20,21}. GTP is required for dopamine stimulation of adenylate cyclase in membranes prepared from homogenates of rat striatum²⁴. It is interesting that the hormonal sensitivity of striatal adenylate cyclase, uncoupled from its receptor by exposure to the venom of Russell's viper, was restored by the addition of GTP to the incubation medium²¹. The most direct evidence that the effect of GTP is related to the coupling of receptors with adenylate cyclase has come from studies of the β -adrenergic receptor–adenylate cyclase system in S49 lymphoma cells. In wild-type S49 cells, guanine nucleotides increase the K_d values and Hill coefficients for agonist, but not antagonist, inhibition of ^{125}I -iodohydroxybenzylpindolol binding²⁰. In contrast, S49 cells, in which the β -adrenergic receptor and adenylate cyclase are both present but 'uncoupled', lack an agonist-specific response to GTP. So far, only receptors that can be shown to be linked to adenylate cyclase have been reported to show decreased affinity for agonists in the presence of guanine nucleotides^{1,5–9}. Opiate receptors, which seem to be linked to adenylate cyclase in some systems²², have been reported to show nucleotide regulation of both agonist and antagonist binding¹⁰.

The relationship of neuroleptic binding sites to dopamine receptors has been questioned. The affinities of dopamine-receptor agonists and antagonists differ depending on whether they are determined in the presence of an agonist or antagonist ligand²³. Discrepancies also exist between the agonist and antagonist potencies determined by binding studies when

Table 1 Effect of GTP on the affinity and cooperativity of agonist and antagonist inhibition of ^3H -spiroperidol binding to rat striatal membranes

	Control	K_d (nM) GTP	Ratio	Hill coefficient Control	GTP
Agonists					
Complete					
Epinephrine	390 $\pm$ 100	1400 $\pm$ 200*	3.6	0.56 $\pm$ 0.05	0.61 $\pm$ 0.02
Dopamine	380 $\pm$ 80	1800 $\pm$ 300*	4.7	0.54 $\pm$ 0.02	0.59 $\pm$ 0.03
(±)ADTN	14 $\pm$ 3	54 $\pm$ 6*	3.9	0.49 $\pm$ 0.07	0.59 $\pm$ 0.07
Partial					
Apomorphine	25 $\pm$ 4	90 $\pm$ 29*	3.6	0.61 $\pm$ 0.04	0.76 $\pm$ 0.06*
Antagonists					
(+)-Butaclamol	0.61 $\pm$ 0.10	0.67 $\pm$ 0.19	1.1	0.65 $\pm$ 0.03	0.66 $\pm$ 0.04
(-)-Butaclamol	3400 $\pm$ 400	3900 $\pm$ 600	1.1	0.97 $\pm$ 0.01	0.94 $\pm$ 0.05
α -Flupenthixol	0.37 $\pm$ 0.04	0.45 $\pm$ 0.13	1.2	0.65 $\pm$ 0.07	0.62 $\pm$ 0.06
β -Flupenthixol	52 $\pm$ 11	52 $\pm$ 4	1.0	1.02 $\pm$ 0.06	0.96 $\pm$ 0.04

Assays were carried out as described in the legend to Fig. 1a. Each ratio of K_d values was calculated as the K_d in the presence of GTP to the K_d in the absence of GTP. The concentration of GTP was 0.3 mM. The results are expressed as the mean $\pm$ s.e.m. calculated from three to four independent determinations.

* Means values significantly different; Student's *t* test, $P < 0.05$.

compared with those determined in studies of dopamine-stimulated adenylate cyclase activity^{13,23}. Further, the subcellular distributions of ^3H -spiroperidol binding sites and adenylate cyclase in the rat brain are not identical²⁴. From such studies, it has been suggested that the dopamine receptor and adenylate cyclase are related but not identical sites. Also, decreases in dopamine-sensitive adenylate cyclase activity (85%) and ^3H -haloperidol binding (40%) occur following intrastriatal injection of kainic acid²⁵. The difference in the magnitude of the effects of kainic acid has led some investigators to question whether striatal adenylate cyclase and ^3H -haloperidol binding sites are associated. The data presented here demonstrate a specific effect of GTP on the inhibition of ^3H -spiroperidol binding by the same agonists which have previously been shown to activate adenylate cyclase^{11,12}. The connection between the striatal ^3H -spiroperidol binding site and the dopamine-sensitive adenylate cyclase is thus strengthened^{15,16}.

We thank Dr Kenneth P. Minneman for suggestions and discussions and Dr John McDermid, Wellcome, for (±)ADTN. N. R. Z. is a postdoctoral fellow of the USPHS (NS 05970). This work was supported by the USPHS (NS 09199, NS 13289 and AA 03537).

NANCY R. ZAHNISER
PERRY B. MOLINOFF

Department of Pharmacology,
University of Colorado Medical Center,
4200 East Ninth Avenue,
Denver, Colorado 80262

Received 5 June; accepted 7 August 1978.

- Rodbell, M. *et al.* *Adv. Cyclic Nucleotide Res.* **5**, 3–29 (1975).
- Clement-Cormier, Y. C., Parrish, R. G., Petzold, G. L., Keabian, J. W. & Greengard, P. J. *J. Neurochem.* **25**, 143–149 (1975).
- Brown, E. M., Fedak, S. A., Woodard, C. J. & Aurbach, G. D. *J. biol. Chem.* **251**, 1239–1246 (1976).
- Ronfagalis, B. D., Thornton, M. & Wade, D. N. *J. Neurochem.* **27**, 1533–1535 (1976).
- Maguire, M. E., Ross, E. M. & Gilman, A. G. *Adv. Cyclic Nucleotide Res.* **8**, 1–83 (1977).
- Rodbell, M., Krans, H. M. J., Pohl, S. L. & Birnbaumer, L. *J. biol. Chem.* **246**, 1872–1876 (1971).
- Maguire, M. E., Van Arsdale, P. M. & Gilman, A. G. *Molec. Pharmac.* **12**, 335–339 (1976).
- Williams, L. T. & Lefkowitz, R. J. *J. biol. Chem.* **252**, 7207–7213 (1977).
- Lefkowitz, R. J., Mullikin, D., Wood, C. L., Gore, T. B. & Mukherjee, C. J. *J. biol. Chem.* **252**, 5295–5303 (1977).
- Blume, A. J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1713–1717 (1978).
- Miller, R. J., Horn, A. S. & Iversen, L. L. *Molec. Pharmac.* **10**, 759–766 (1974).
- Iversen, L. L. *Science* **188**, 1084–1089 (1975).
- Goldberg, L. I., Volkman, P. H. & Kohli, J. D. A. *Rev. Pharmac. Toxic.* **18**, 57–79 (1978).
- Creese, I., Schneider, R. & Snyder, S. H. *J. Pharmac.* **46**, 377–381 (1977).
- Fields, J. Z., Reisine, T. D. & Yamamura, H. I. *Brain Res.* **136**, 578–584 (1977).
- Leysen, J. E., Gommeren, W. & Laduron, P. M. *Biochem. Pharmac.* **27**, 307–316 (1978).
- Seeman, P., Chau-Wong, M., Tedesco, J. & Wong, K. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4376–4380 (1975).
- Burt, D. R., Enna, S. J., Creese, I. & Snyder, S. H. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4655–4659 (1975).
- Roberts, P. J., Woodruff, G. N. & Poat, J. A. *Molec. Pharmac.* **13**, 541–547 (1977).

- Haga, R., Ross, E., Anderson, H. J. & Gilman, A. G. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2016–2020 (1977).
- Keabian, J. *Brain Res.* **144**, 194–198 (1978).
- Klee, W. A. & Nirenberg, M. *Nature* **263**, 609–612 (1976).
- Burt, D. R., Creese, I. & Snyder, S. H. *Molec. Pharmac.* **12**, 800–812 (1976).
- Laduron, P. M., Janssen, P. F. M. & Leysen, J. E. *Biochem. Pharmac.* **27**, 323–328 (1978).
- Schwarz, R., Creese, I., Coyle, J. T. & Snyder, S. H. *Nature* **271**, 766–768 (1978).

Mammary tumour virus and casein gene transcription during mouse mammary development

MORPHOGENESIS and differentiation of the murine mammary gland is dependent on the interaction of peptide and steroid hormones. These hormones cause mammary growth through proliferation of ductal and alveolar epithelial cells^{1,2} and differentiation of the gland for the production of the milk proteins, casein and α -lactalbumin^{2–4}. Mammary development frequently increases the incidence and results in an earlier onset of mammary tumours^{5,6}. The influence of mammary development upon tumorigenesis has been correlated with the appearance of mouse mammary tumour virus (MMTV) particles^{6–10}. Importantly, the production of MMTV particles is dependent on alveolar differentiation and function^{7,11}. Lactating mammary glands from BALB/c mice contained considerably lower levels of MMTV RNA than glands from C3H mice infected with the milk-transmitted virus MMTV-S (ref. 12). In this study we have examined the influence of the hormonal changes that occur during mammary development on MMTV and casein expression by molecular hybridisation with specific complementary DNA (cDNA) probes to MMTV 70S RNA and casein 15S mRNA. Hence, the relationship between MMTV expression and the expression of a normal differentiated mammary function—the production of casein mRNA—was determined. We have quantitated MMTV RNA and casein mRNA in developing mammary tissues from BALB/c and BALB/cfC3H mice that have a low and high incidence of mammary tumours respectively^{6,13}. These studies demonstrate that the hormonal alterations occurring during mammary development enhanced MMTV RNA expression to a greater extent in the high than the low tumour-incidence strain, whereas similar changes in the level of casein mRNA were observed in both strains during mammary development.

A representative MMTV cDNA probe was synthesised *in vitro* by avian myeloblastosis virus reverse transcriptase using 70S MMTV RNA as template and calf thymus DNA fragments¹⁴ as a random primer^{15,16}. The level of MMTV RNA was quantitated in cellular RNA extracts by the titration method¹⁷. The amount of MMTV RNA, expressed as a percentage of the

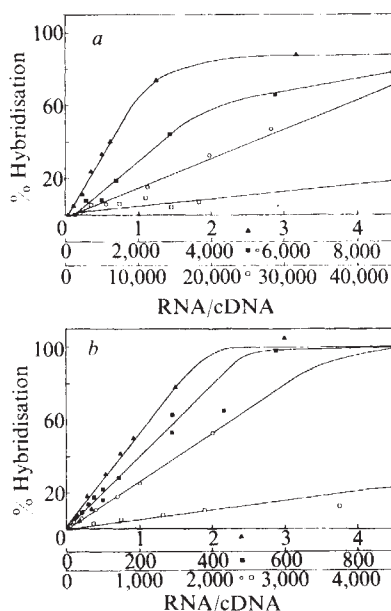


Fig. 1 Detection of MMTV RNA and mouse casein 15S mRNA in mammary glands by molecular hybridisation. Cellular RNA was isolated by sodium dodecyl sulphate-phenol extraction with proteinase K digestion²⁰ followed by extraction with 3 M sodium acetate²¹. Hybridisations were carried out and assayed by S1 nuclease digestion¹⁹. *a*, Hybridisations of MMTV cDNA, equivalent Dot of 0.09, with MMTV 70S RNA isolated from purified MMTV virions (▲) and RNA isolated from early pregnant (□), late pregnant (■) and late lactating (○) uniparous BALB/cfC3H mammary tissues. *b*, Hybridisations of mouse casein cDNA, equivalent Dot of 0.09, with purified mouse casein 15S mRNA (▲) and RNA isolated from early (□), mid (○) and late (■) pregnant uniparous BALB/cfC3H mammary tissues.

cellular RNA, was determined by comparing the rate of hybridisation observed with each cellular RNA to the rate of hybridisation observed with purified 70S MMTV RNA (Fig. 1*a*). Casein mRNA was quantitated in the same manner (Fig. 1*b*) with a full-length cDNA probe synthesised from purified mouse casein 15S mRNA¹⁸ as described previously¹⁹. The limit of detection of MMTV RNA and casein mRNA was 0.0005% of the total cellular RNA.

Casein mRNA was quantitated in mammary tissues from virgin and uniparous BALB/c and BALB/cfC3H mice at various stages of mammary development. Figure 2*a* shows that the level of casein mRNA, which was at a very low level in the virgin gland (0.0009%), was present at an appreciable level in early pregnancy (0.0066%) and then increased to 0.365% of the RNA late in pregnancy. Casein mRNA increased nearly four-fold from late pregnancy to early lactation and was maintained at a high level, approximately 1% of the RNA, during lactation. Following weaning, casein mRNA decreased to one tenth or less of the lactating level, but was considerably higher than observed in the virgin gland. Essentially identical changes in the level of casein mRNA were observed during mammary development in both the low and the high mammary tumour incidence strains, BALB/c and BALB/cfC3H, respectively.

MMTV RNA was quantitated in mammary tissues from virgin and uniparous BALB/c and BALB/cfC3H mice at various stages of mammary development (Fig. 2*b*). The increased tumour incidence of BALB/cfC3H mice relative to the isogenic strain, BALB/c, is due to the milk-transmitted virus, MMTV-S, obtained by foster-nursing BALB/c mice on C3H. Note that normal BALB/c and BALB/cfC3H tissues contain 5–9 copies of MMTV DNA per diploid cell²². The level of MMTV RNA in the mammary tissues was below the limit of detection in virgin and early pregnant BALB/cfC3H mammary glands and was present at low levels (0.0014%) in mid pregnancy. MMTV RNA was at an appreciably higher level (0.0125%) in the

BALB/cfC3H mammary gland at late pregnancy and increased approximately twofold after parturition. Later in lactation, the level of MMTV RNA decreased from the level observed in early lactating BALB/cfC3H mammary glands. The decrease in MMTV RNA late in lactation may be due to the release of MMTV RNA particles into glandular secretions^{6–8,10,11}. MMTV RNA levels were similar in BALB/cfC3H mammary glands late in lactation and 10 days after weaning, 0.0076% and 0.0096%, respectively. Large variations in the amount of MMTV RNA (Fig. 2*b*) but not casein mRNA (Fig. 2*a*) were observed in pregnant and lactating BALB/cfC3H mammary tissues. Interestingly, differences in milk virus titre, particularly in uniparous mice, have been reported to be indicative of similar biological variation in virus production⁸. The level of MMTV RNA in BALB/c mammary tissues at all the developmental stages assayed was 0.0005% or less.

The hormonal changes that control normal mammary development also influence neoplastic development of the mammary gland^{5,6}. In BALB/c and BALB/cfC3H mice the mammary tumour incidence is higher in retired breeders than virgin mice (refs 6, 13 and D. Medina, personal communication). The occurrence of MMTV antigens in milk from lactating mice increases with the number of lactations in certain mouse strains^{9,10,23}. Mammary tumours from BALB/c and BALB/cfC3H retired breeders have higher levels of MMTV RNA than observed in uniparous lactating mammary glands (R. J. P. and S. H. S., in preparation). These observations suggest that additional cycles of hormonal stimulation during pregnancy and lactation may increase the level of MMTV RNA in mammary glands. Therefore, the level of MMTV RNA was quantitated in mammary glands from BALB/c and BALB/cfC3H lactating multiparous mice and retired breeders (Table 1). The levels of casein mRNA in multiparous lactating mice and retired

Fig. 2 Mouse casein 15S mRNA and MMTV RNA in mammary tissues during mammary development. Tissues were obtained from BALB/c and BALB/cfC3H virgin mice (Vi) and either pregnant (Pg), lactating (La) or post-weaning (PW) uniparous mice. Early (E), mid (M) and late (L) pregnant mice were staged at 1–7, 8–14, and 15–21 days of pregnancy, respectively. Early (E), mid (M) and late (L) lactating mice were staged at 1–7, 8–14 and 15–21 days of lactation respectively. Post-weaning mammary glands were obtained from uniparous mice 5 or 10 days following weaning. Cellular RNA was prepared and hybridisations were carried out as described in the legend to Fig. 1. The amount of casein mRNA (Fig. 2*a*) and MMTV RNA (Fig. 2*b*) are expressed as a percentage of the cellular RNA. Solid bars denote BALB/c tissues and cross-hatched bars denote BALB/cfC3H tissues, presented as the mean $\pm$ one s.e.m. for (*n*) determinations.

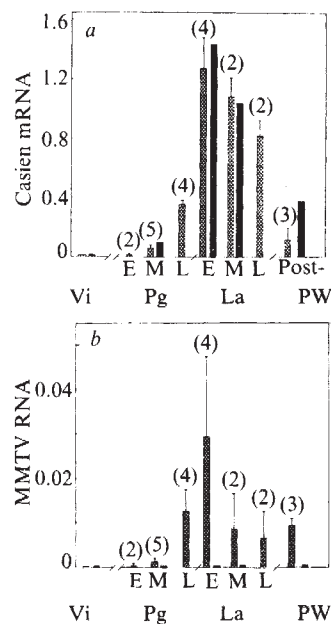


Table 1 MMTV RNA and casein mRNA in mammary tissue from multiparous mice and retired breeders

Strain	Developmental stage	MMTV RNA (% total RNA)	Casein mRNA (% total RNA)
BALB/cfC3H	Lactating-multiparous	0.0181	1.700
BALB/cfC3H	Retired breeder	0.0128	0.017
BALB/c	Lactating-multiparous	0.0006	1.790
BALB/c	Retired breeder	0.0009	0.014

Mammary tissue excised from 5–7 day lactating multiparous mice and retired breeders, obtained one month after weaning of the fourth litter, were used for the assays. Cellular RNA was prepared and hybridisations were performed as described in the legend to Fig. 1.

breeders of both strains were consistent with the developmental stage of the gland. In BALB/cfC3H mammary glands from multiparous lactating mice and retired breeders (Table 1), the level of MMTV RNA was similar to that observed in uniparous lactating, 0.0295%, and weaned, 0.0096%, mice. Mammary tissues from BALB/c retired breeders contained a low but detectable amount of MMTV RNA (0.0009%) whereas uniparous and multiparous BALB/c lactating mammary tissues contained lower levels of MMTV RNA ($\leq 0.0006\%$). Varmus *et al.*¹² also have reported that lactating BALB/c mammary glands contained low levels of MMTV RNA (0.0003%). Therefore, additional cycles of hormonal stimulation increased MMTV expression in BALB/c but not BALB/cfC3H mammary glands.

These studies indicate that the hormonal milieu required for mammary development and lactogenesis enhanced both casein and MMTV expression in the mammary gland of the high tumour incidence strain, BALB/cfC3H. Both MMTV RNA and casein mRNA were detectable by mid pregnancy and increased until the maximum level was observed early in lactation. Although the level of both MMTV RNA and casein mRNA decreased following weaning in both uniparous and multiparous mice, MMTV RNA did not decrease to the same degree as casein mRNA. Therefore, the accumulation of MMTV RNA in BALB/cfC3H mammary glands is dependent upon mammary development and influenced by differentiated mammary function. Consequently, the accumulation of B-type virus particles, which is dependent upon alveolar differentiation and function^{8,11}, is also related to and may be dependent upon accumulation of MMTV RNA in the mammary gland of high tumour incidence strains. Interestingly, an appreciable level of MMTV RNA was maintained in the absence of continued hormonal stimulation of the BALB/cfC3H gland. Casein expression in the low tumour incidence strain, BALB/c, was influenced similarly by the hormonal changes that occur during mammary development. However, MMTV expression during mammary development in BALB/c mice was not influenced to the same extent as observed in BALB/cfC3H mice. Therefore, MMTV expression in the low tumour incidence strain is apparently not influenced by hormonally-regulated mammary development to the same degree as in the high tumour incidence strain.

Genetic and biochemical evidence indicate that mouse strains contain different variants of MMTV^{6,24–28}. MMTV-O is apparently an endogenous virus transmitted as a germinal provirus in BALB/c mice^{24,25}, while BALB/cfC3H mice contain both MMTV-O and the exogenous virus, MMTV-S, transmitted by foster-nursing. Our studies indicated that MMTV-O and MMTV-S do not respond to the same extent to changes occurring in the hormonal environment during normal mammary development. Further studies using specific nucleic acid probes²⁶ or antibodies to type-specific antigenic determinants^{27,28} should provide additional evidence that MMTV-O and MMTV-S respond differently to changes in the hormonal environment during normal mammary development. The effect of mammary development on MMTV expression is related to

the tumour incidence of the strain, suggesting that multiple factors may influence MMTV expression during normal development.

This research was supported by contract NO1-CP-4-3385 within the NCI Virus Cancer Program. Purified MMTV virions and avian myeloblastosis virus reverse transcriptase were obtained through the Office of Resources and Logistics of the Virus Cancer Program. We thank Owen DeWitt for technical assistance and Dr Daniel Medina for providing the mice and reviewing the manuscript.

ROBERT J. PAULEY
JEFFREY M. ROSEN
SUSAN H. SOCHER

Department of Cell Biology,
Baylor College of Medicine,
Houston, Texas 77030

Received 22 May; accepted 31 July 1978.

- Ceriani, R. L. *J. invest. Derm.* **63**, 93–108 (1974).
- Banerjee, M. R. *Int. Rev. Cytology* **47**, 1–97 (1976).
- Topper, Y. J. *Rec. Prog. Horm. Res.* **26**, 287–308 (1970).
- Turkington, R. W. *Biochem. Actions Horm.* **2**, 55–80 (1971).
- Lathrop, A. E. C. & Leob, L. *Proc. Soc. exp. Biol. Med.* **11**, 38–40 (1913).
- Nandi, S. & McGrath, C. M. *Adv. Cancer Res.* **17**, 353–414 (1973).
- Feldman, D. G. *J. natn. Cancer Inst.* **30**, 477–502 (1963).
- Huesby, R. A., Barnum, C. P. & Bittner, J. J. *Cancer Res.* **10**, 516–520 (1950).
- Charney, J., Pollinger, B. D. & Moore, D. H. *J. natn. Cancer Inst.* **43**, 1289–1296 (1969).
- Nowinski, R. C., Old, L. J., Moore, D. H., Geering, G. & Boyse, E. A. *Virology* **31**, 1–14 (1967).
- Lasfargues, E. Y. & Feldman, D. G. *Cancer Res.* **23**, 191–201 (1963).
- Varmus, H. E. *et al. J. molec. Biol.* **79**, 663–679 (1973).
- Medina, D., DeOme, K. B. & Young, L. J. *J. natn. Cancer Inst.* **44**, 167–174 (1970).
- Taylor, J. M., Illmensee, R. & Summers, J. *Biochim. biophys. Acta* **442**, 324–330 (1976).
- Kacian, D. L. & Myers, J. C. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2191–2195 (1976).
- Dudley, J., Butel, J. S., Socher, S. H. & Rosen, J. M. *J. Virol.* (submitted).
- Young, B. D. *et al. J. molec. Biol.* **84**, 555–568 (1974).
- Rosen, J. M. *Biochemistry* **15**, 5263–5271 (1976).
- Rosen, J. M. & Barker, S. W. *Biochemistry* **15**, 5272–5280 (1977).
- Rosen, J. M. & Socher, S. H. *Nature* **269**, 83–86 (1977).
- Palmiter, R. J. *J. biol. Chem.* **248**, 2095–2106 (1973).
- Morris, V. L., Medeiros, E., Ringold, G. M., Bishop, J. M. & Varmus, H. E. *J. molec. Biol.* **114**, 73–91 (1977).
- Morre, D. H., Charney, J. & Pullinger, B. D. *J. natn. Cancer Inst.* **45**, 561–565 (1970).
- Bentvelzen, P., Daams, J. H., Hageman, P. & Calafat, J. *Proc. natn. Acad. Sci. U.S.A.* **67**, 377–384 (1970).
- Bentvelzen, P. *Biochim. biophys. Acta* **355**, 236–259 (1974).
- Drohan, W., Kettmann, R., Colcher, D. & Schlom, J. *J. Virol.* **21**, 986–995 (1977).
- Teramoto, Y. A., Kufe, D. & Schlom, J. *J. Virol.* **24**, 525–533 (1977).
- Teramoto, Y. A. & Schlom, J. *Cancer Res.* **38**, 1990–1995 (1978).

Male recombination is not induced in *Drosophila melanogaster* by extracts of strains with male recombination potential

NON-TRIVIAL frequencies of spontaneous male recombination in the hybrid progeny of *Drosophila melanogaster* males from so-called 'male recombination strains' and laboratory marker stock females are well documented^{1,2}. Attempts to induce male recombination by injection have produced conflicting results. The positive results originally claimed by Reddi *et al.*³ were not reproducible in two subsequent studies^{4,5}. Sochacka and Woodruff⁶ have suggested that the disparity in results may be attributable to the use of extracts from a male recombination strain in the former study only. They injected extracts from the male recombination strain OK1 (ref. 7) into laboratory marker strain females. These were crossed to the Canton-S strain, which lacks spontaneous male recombination potential, producing F₁ hybrids which yielded a low frequency of male recombination. Injection directly into F₁ hybrid males also yielded some recombinants. Similar tests using Canton-S as the source of the extract were negative. We describe here large-scale tests with different male recombination strains. We could not reproduce the injection effect of Sochacka and Woodruff, suggesting that their results may not be generally applicable. We suggest, furthermore, that the genetic implications of a negative injection result may be at least as interesting as those of a positive result. A positive result would imply the existence of some co-existing,

Table 1 Observed numbers of recombinants and the minimum number of recombination events in various studies

	Experimental		Total	Control		Total
	Actual	Minimum events		Actual	Minimum events	
Sydney	7	6	116,776	9	6	70,548
Brown	9	6	39,628	22	15	86,024
Sochacka and Woodruff ⁶	25	11	23,170	0	0	67,910

'Experimental' includes all cases of injections of extracts from male recombination strains into females and males.

'Control' includes injection with non-recombination strains, injection of solvents and no injection.

possibly non-essential, factor, such as a virus or episome. This would rule out explanations for male recombination based on incompatibility of genetic organisation in different strains. Explanations of this type, for example, in terms of spatial organisation of chromosomes⁸, have potentially important implications for normal nuclear organisation.

The experiments were carried out independently at Sydney University and at Brown University. Extracts from the strong male recombination strains, Harwich² and Hunter Valley⁸, were used in the experimental groups and extracts from the laboratory strain Canton-S were used as controls in both laboratories. In Sydney, Harwich or Hunter Valley extracts were injected into females of the laboratory strain *al cn bw* (chromosome II). At Brown, Harwich extracts were injected into the laboratory strains *al cl b c sp²* (chromosome II) and *rucuca* (chromosome III). The majority of control flies were injected with Canton-S extracts, a few were injected with Ringer's solution and a few were untreated. Experimental and control tubes were coded and combined for counting, an important consideration where low recombination frequencies are expected.

Altogether, 10 experiments were carried out. The initial experiments were made using the concentration of extract given by Sochacka and Woodruff⁶. Subsequently, concentrations up to four times higher were used. We also examined the effect of ageing the injected flies, in case a specific developmental stage or incubation period was important. In addition, following a suggestion by Dr D. Angus of Newcastle University, New South Wales, we injected extracts from various hybrids rather than from pure strains, on the grounds that male recombination might be attributable to an agent which multiplies preferentially in the hybrid.

None of the experiments showed an elevated level of recombination. Results have therefore been pooled within laboratories, and are presented in the first two lines of Table 1. We show both the actual number of recombinants found and the minimum number of independent recombination events⁹. Neither of these numbers is an entirely satisfactory measure of recombination intensity. The actual number of recombinants cannot validly be used in statistical tests if recombinants occur in clusters rather than independently of each other. On the other hand, comparisons based on the minimum number of recombination events are strictly valid only if a fixed number of progeny is screened per cross. In the present experiments, however, neither measure indicates an effect of the treatment.

The data of Sochacka and Woodruff⁶ are summarised in the third line of Table 1. The complete absence of recombinants in their control series is noteworthy considering that Canton-S was used by all three laboratories. It is a major cause of the significance of the difference in recombination frequency between their experimental and control groups. Also, the frequency in the experimental injected group is about an order of magnitude lower than the spontaneous frequency in OK1 hybrid males⁷.

There are, of course, several possible reasons for dissimilar results using different strain combinations. We suggest, nevertheless, that the hypothesis of direct transmissibility of male recombination by injection requires further substantiation.

This work was supported in part by the Australia Research Grants Commission and by NSF grant DEB76-82630. We

thank Dr Anne Fausto-Sterling for technical advice and the loan of equipment.

J. A. SVED
D. C. MURRAY

*School of Biological Sciences,
University of Sydney, New South Wales, Australia*

R. E. SCHAEFER
M. G. KIDWELL*

*Division of Biology and Medicine,
Brown University, Providence, Rhode Island 02912*

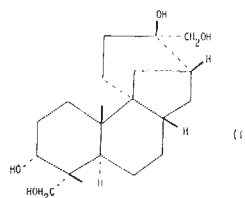
Received 17 May; accepted 11 August 1978.

* To whom reprint requests should be addressed.

1. Hiraizumi, Y. *Proc. natn. Acad. Sci. U.S.A.* **68**, 268-270 (1971).
2. Kidwell, M. G., Kidwell, J. F. & Sved, J. A. *Genetics* **86**, 813-833 (1977).
3. Reddi, O. S., Reddy, G. M. & Rao, M. S. *Nature* **208**, 203 (1965).
4. Singer, K. M., Chovnick, A., Suzuki, D. T., Baillie, D. & Hoar, D. *Nature* **214**, 503-504 (1967).
5. Saraswathy, T. *Experientia* **26**, 799 (1970).
6. Sochacka, J. H. M. & Woodruff, R. C. *Nature* **262**, 287-289 (1976).
7. Woodruff, R. C. & Thompson, J. N. *Heredity* **38**, 291-308 (1977).
8. Sved, J. A. *Aust. J. Biol. Sci.* **29**, 375-388 (1976).
9. Kidwell, M. G. & Kidwell, J. F. *J. Hered.* **66**, 367-375 (1975).

Aphidicolin prevents mitotic cell division by interfering with the activity of DNA polymerase- α

THE study of the control of DNA synthesis would be helped if a variety of chemicals which influenced DNA synthesis in various ways was available. We have searched for chemicals which prevent mitotic division of sea urchin embryos, which requires DNA synthesis, but do not prevent meiotic divisions of starfish oocytes, which are independent of DNA synthesis¹. Aphidicolin² (I) is one such compound. Furthermore, recent experiments have shown that aphidicolin selectively inhibits the activity of DNA polymerase- α obtained from regenerating rat liver without interfering with the activity of DNA polymerase- β and mitochondrial DNA polymerase³. The function of the DNA polymerases in the complex process of DNA replication is not known. Correlative studies of DNA synthesis and the level of DNA polymerase activity suggest that DNA polymerase- α may be important in DNA replication⁴. However, in certain cases, DNA polymerase- β and - γ also increase at the time of DNA synthesis, and attempts to specify which of the DNA polymerases is 'replicative' or to assign roles to the various polymerases have not yet been successful⁴. We report here evidence that prevention of cell division by aphidicolin in sea urchin embryos is due to selective inhibition of DNA polymerase- α activity. This study implies a functional role for this polymerase activity during replication of DNA.



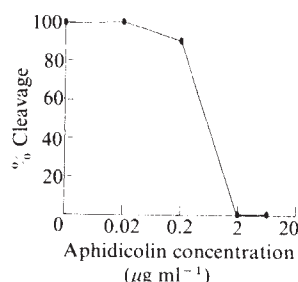


Fig. 1 Inhibition by aphidicolin of mitotic division of embryos of *Hemicentrotus pulcherrimus*. Five min after fertilisation, embryos were treated with various concentrations of aphidicolin. Morphological changes were observed for 150 min. Aphidicolin (insoluble in aqueous media) was dissolved in DMSO and then dispersed in seawater to give a DMSO concentration of 0.1% (v/v) in the solution, which did not affect cleavage.

When embryos of the sea urchin *Hemicentrotus pulcherrimus* were exposed to aphidicolin ($2 \mu\text{g ml}^{-1}$) for 150 min from 5 min after insemination, nearly all of them remained at the 1-cell stage whereas control embryos divided three times (Fig. 1). The incorporation of ^3H -thymidine into DNA of the embryos was inhibited by 90% by treatment with aphidicolin at $2 \mu\text{g ml}^{-1}$ (Table 1). The rate of protein synthesis was the same whether or not embryos were exposed to the compound (Table 1). RNA synthesis remains quite low during early cleavage and maternal messenger RNA synthesised during oogenesis and stored in eggs plays a central role in translation⁵. Therefore, the effect of aphidicolin on RNA synthesis was examined at the gastrula stage. The incorporation of ^{14}C -uridine into the acid-insoluble material was not affected by 1 h exposure to aphidicolin ($10 \mu\text{g ml}^{-1}$).

Aphidicolin ($40 \mu\text{g ml}^{-1}$) did not prevent 1-methyladenine-induced (45 ng ml^{-1}) meiotic maturational divisions in oocytes of the starfish *Asterias amurensis*⁶. Meiosis is a reduction in chromosome number and no chromosomal replication takes place during the maturational divisions. Thus, it seems that aphidicolin prevents the sequence of events necessary to complete chromosomal replication in sea urchin embryos rather than the process of karyokinesis and cytokinesis.

Inhibitors of deoxyribonucleotide biosynthesis such as *N*-hydroxyurea ($200 \mu\text{g ml}^{-1}$) and 5-fluorouracil ($200 \mu\text{g ml}^{-1}$) did not affect the first cleavage of sea urchin embryos. In fact,

Table 1 Effects of aphidicolin on DNA and protein synthesis in embryos of *Hemicentrotus pulcherrimus*

Aphidicolin concentration ($\mu\text{g ml}^{-1}$)	^3H -Thymidine incorporation (%)	^3H -L-Leucine incorporation (%)
0	100	100
0.02	76	105
0.2	38	125
2	10	124
10	5	110

Five min after fertilisation, ^3H -thymidine ($0.63 \mu\text{Ci}$; 24 Ci mmol^{-1}) and various amounts of aphidicolin were added to 5×10^5 eggs in 5 ml of seawater. Ninety min later, embryos were sedimented, washed, fixed and extracted with 1 M perchloric acid, and the acid-insoluble radioactivity was counted according to the method of Hinegardner *et al.*⁸. The radioactivity in the control experiment was 1.7×10^4 c.p.m. ^3H -L-Leucine (0.2 mCi ; 58 Ci mmol^{-1}) was added to 4×10^5 unfertilised eggs suspended in 20 ml of seawater. Five min later, eggs were thoroughly washed with seawater, suspended in 20 ml of seawater, inseminated and various amounts of aphidicolin were added to 2.5-ml aliquots of the suspension. Ninety min later, embryos were collected, fixed and extracted with 10% TCA at 90°C , and the acid-insoluble radioactivity was counted according to the method of Mano⁹. The control value was 5.9×10^5 c.p.m. Aphidicolin was dissolved in DMSO and then dispersed in seawater to give a final concentration of 0.1% DMSO (v/v), which did not affect DNA or protein synthesis in the embryo.

Table 2 Effects of aphidicolin on activity of DNA polymerase- α , - β and - γ obtained from the nuclear fraction of blastulae of *Hemicentrotus pulcherrimus*

Aphidicolin concentration ($\mu\text{g ml}^{-1}$)	DNA polymerase activity (%)		
	α	β	γ
0	100	100	100
0.5	39	—	—
2	10	102	95
10	0	99	98

DNA polymerase- α was purified from the nuclear extract using phase separation and hydroxyapatite and phosphocellulose columns. The properties of the polymerase were similar to those described by Loeb¹⁰. DNA polymerase- β was obtained as described elsewhere¹¹. The activated DNA-dependent DNA synthesis with DNA polymerase- α and - β were determined using 50 mM Tris-maleate buffer, pH 8.5, as described previously¹¹. The poly(rA)-oligo(dT)₁₂₋₁₈ dependent DNA polymerase- β and - γ activities were measured as described elsewhere (Habara, A., Nagano, H. & Mano, Y., in preparation). In the case of DNA polymerase- β , the values shown in this table were obtained by using activated DNA as template primer. The poly(rA)-oligo(dT)₁₂₋₁₈ dependent DNA polymerase- β activity was also insensitive to aphidicolin. The activity of DNA polymerase- α , - β and - γ in the absence of aphidicolin was 3,200, 600 and 5,200 c.p.m., respectively. Aphidicolin was dissolved in DMSO and then dispersed in the reaction mixtures to give a final DMSO concentration of less than 0.1% (v/v). At these concentrations, DMSO did not affect the activity of the DNA polymerases.

unfertilised sea urchin eggs have enough pools of four deoxyribonucleoside triphosphates to support two or three cycles of DNA replication⁷. This excludes the possibility that aphidicolin acts by inhibiting deoxyribonucleotide biosynthesis, thereby depriving DNA polymerases of necessary substrates.

The effects of aphidicolin on the activity of DNA polymerase- α , - β and - γ obtained from the nuclear fraction of the sea urchin embryos (blastulae) were next examined. The addition of aphidicolin ($2 \mu\text{g ml}^{-1}$) inhibited the activity of DNA polymerase- α by 90% whereas that of DNA polymerase- β and - γ was not affected by the treatment (Table 2). Thus the degree of inhibition by aphidicolin of the activity of DNA polymerase- α was nearly the same as that observed for cleavage and DNA synthesis in the intact cell (Fig. 1, Tables 1 and 2). However, it was observed that a higher dose of aphidicolin was required to inhibit the activity of DNA polymerase- α which was further purified by a DEAE-Sephadex A-50 column; aphidicolin at $2 \mu\text{g ml}^{-1}$ produced 58% inhibition of DNA synthesis with this preparation. This could be due to a change in the enzyme molecule from its native state during purification or could be related to the adherence of some component to the enzyme which facilitates the action of the drug and is separated from the enzyme by DEAE-Sephadex chromatography. We are at present carrying out further studies designed to test these possibilities. Similarly, aphidicolin inhibited the activity of DNA polymerase- α obtained from unfertilised eggs but did not affect DNA polymerase- β activity. These results suggest that DNA polymerase- α plays an essential part in chromosomal DNA replication in cleavage of sea urchin embryos. Thus aphidicolin has considerable potential for investigating the fundamental nature of the cell functions concerned.

We thank A. Habara for a supply of DNA polymerase- γ . This work was partly supported by a grant from the Ministry of Education, Science and Culture, Japan.

SUSUMU Ikegami

Department of Agricultural Chemistry,
University of Tokyo

TAKAHIKO Taguchi
MOCHIIKO Ohashi

Department of Biochemistry,
Tokyo Metropolitan Institute of Gerontology,
Itabashi-ku, Tokyo 173

MIEKO OGURO
HIROSHI NAGANO
YOSHITAKE MANO

Department of Physiological Chemistry and Nutrition,
Faculty of Medicine,
University of Tokyo,
Bunkyo-ku, Tokyo 113, Japan

Received 1 June; accepted 4 August 1978.

1. Zampetti-Bosseler, F., Huez, G. & Brachet, J. *Expl Cell Res.* **78**, 383–393 (1973).
2. Dalziel, W., Hesp, B., Stevenson, K. M. & Jarvis, J. A. *J. chem. Soc. Perkin Trans. I*, 2841–2851 (1973).
3. Ohashi, M., Taguchi, T. & Ikegami, S. *Biochem. biophys. Res. Commun.* **82**, 1084–1090 (1978).
4. Weissbach, A. A. *Rev. Biochem.* **46**, 25–47 (1977).
5. Brachet, J. & Malpoix, P. *Adv. Morphogenesis* **9**, 263–316 (1971).
6. Ikegami, S., Kamiya, Y. & Shirai, H. *Expl Cell Res.* **103**, 233–239 (1976).
7. Mathews, C. K. *Expl Cell Res.* **92**, 47–56 (1975).
8. Hinegardner, R. T., Rao, B. & Feldman, D. E. *Expl Cell Res.* **36**, 51–61 (1964).
9. Mano, Y. *Dev Biol.* **22**, 433–460 (1970).
10. Loeb, L. A. *J. biol. Chem.* **244**, 1672–1681 (1969).
11. Suzuki-Hori, C., Nagano, H. & Mano, Y. *J. Biochem., Tokyo* **82**, 1613–1621 (1977).

Resistance to viomycin conferred by RNA of either ribosomal subunit

THE basic peptide antibiotic viomycin inhibits prokaryotic protein biosynthesis^{1,2}, and seems to bind at least to the small ribosomal subunit because it competes partially with the binding of streptomycin³. However, viomycin does not induce misreading like streptomycin¹, but promotes association of the ribosomal subunits, stabilises 70S couples⁴, and blocks translocation^{5,6}. Interestingly, resistant mutants have been found containing either an altered 30S or an altered 50S subunit^{7–9}. Viomycin strengthens the association of the sensitive subunits but does not affect the association of subunits derived from resistant mutants⁴. These mutants have been obtained from *Mycobacterium smegmatis*, which is highly sensitive in culture in contrast to *Escherichia coli*. We describe here the formation of totally reconstituted subunits using components of sensitive and resistant ribosomes. The reconstituted particles were tested with their complementary, sensitive subunit with respect to viomycin-induced 70S couple formation. The resistance property of either subunit was shown to be located in the RNA moiety.

In the first series of experiments two Rabinowitchi (R) strains of *M. smegmatis* were used: R15 as the viomycin sensitive parental strain and the resistant mutant R33 which has an altered 30S subunit⁹. Intact 16S RNA from 30S subunits of both strains was isolated by the phenol method¹⁰. The isolation of the total proteins (TP30) and the reconstitution procedure for the small subunit have been reported¹¹. RNA and protein moieties from the sensitive and resistant strains were reconstituted in all combinations. The reconstituted particles were supplemented with sensitive 50S (R15) subunits and analysed on sucrose gradients. The conditions of the sucrose gradient centrifugation were chosen so that sensitive ribosomes only formed appreciable amounts of 70S couples in the presence of the drug⁴. Thus, when the ribosomal component conferring the resistance was used for reconstitution, viomycin should not be able to induce the formation of 70S couples.

Figure 1 shows the results. The homologous reconstitutions behaved as expected. 30S particles reconstituted with components of the sensitive strain form 70S couples in the presence of viomycin (Fig. 1b), whereas those obtained with resistant components do not (Fig. 1d). The heterologous reconstitution shows a sensitive response (70S couple formation) when parental 16S RNA was used (Fig. 1f). In contrast, a resistant response was found in the case where mutant 16S RNA was reconstituted with parental TP30. Thus, the resistance

follows the 16S RNA and is not related to the protein moiety (TP30). We conclude that the resistance of the strain R33, which has an altered 30S subunit, is conferred by the 16S RNA.

A similar analysis was performed with native, sensitive 30S subunits and reconstituted 50S particles. 50S subunits were obtained from R15 as the sensitive strain and from the resistant mutant R31 which has an altered 50S subunit⁹. The procedures for the isolation of (23S+5S) RNA and total proteins (TP50)

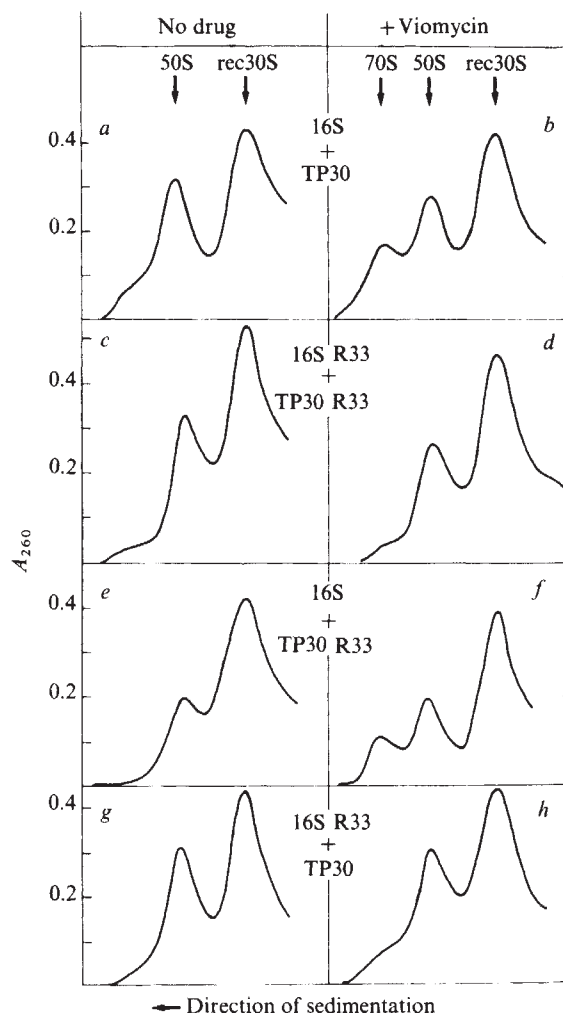


Fig. 1 Sedimentation profiles of reconstituted 30S particles and wild type 50S subunits in the presence and absence of viomycin. In the insets the ribosomal components are listed which were used for reconstitution of the 30S particle: where not indicated as R33 the 16S RNA (16S) and the proteins (TP30) were derived from 30S subunits of the sensitive parental strain R15. Cells were grown and 70S ribosomes prepared as described previously⁷. To remove attached RNase, ribosomes were washed with a buffer containing 10 mM Tris-HCl, pH 7.5, 0.5 M LiCl and 10 mM Mg acetate (ref. 10). The isolation of subunits and RNA (from the subunits which contain intact RNA) was as described¹⁰, the preparation of total proteins TP30 followed ref. 11. For total reconstitution of 30S particles 2.5 A_{260} units of 16S RNA and 0.4 A_{230} units of TP30 in 20 mM Tris-HCl (pH 7.6), 400 mM NH_4Cl , 20 mM Mg acetate, 0.2 mM EDTA and 2 mM β -mercaptoethanol were incubated in 100 μl at 40° for 20 min. One A_{260} unit of reconstituted 30S particles were mixed with 0.4 A_{260} units of parental 50S subunits in a buffer containing 10 mM Tris-HCl, 4 mM Mg acetate and 60 mM NH_4Cl , and where indicated 15 μM viomycin. The mixture was incubated for 10 min at 37° and loaded on a sucrose gradient (10–30%) with the same ionic conditions and where indicated with 15 μM viomycin. After centrifugation for 19 h at 17,000 r.p.m. and 4°C in a Beckman SW40 rotor the gradients were automatically monitored for A_{260} .

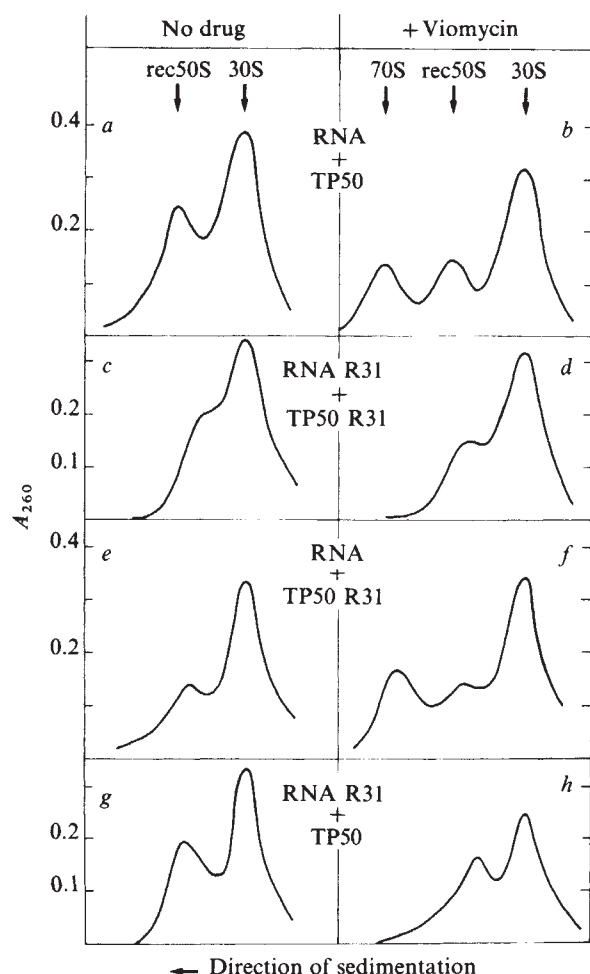


Fig. 2 Sedimentation profiles of wild type 30S subunits and reconstituted 50S particles in the presence and absence of viomycin. Where not indicated as R31 the (23S+5S) RNA (RNA) and the total proteins (TP50) were derived from 50S subunits of the sensitive parental strain R15. Isolation of subunits and RNA was as described in Fig. 1, the preparation of TP50 followed ref. 10. For total reconstitution of 50S particles 2.5 A_{260} units of (23S+5S) RNA and 0.4 A_{230} units of TP50 were incubated in the conditions of the two-step incubation¹⁰. Parental 30S subunits (0.6 A_{260} units) were mixed with 0.4 A_{260} units of reconstituted 50S particles, incubated and analysed on sucrose gradients as described in Fig. 1.

and for the reconstitution of the large subunit have been described¹⁰. Both homologous and heterologous reconstitutions led to particles sedimenting at the 50S position (Fig. 2). In the presence of viomycin we found a pattern equivalent to that shown for reconstituted 30S particles in Fig. 1: All reconstituted 50S particles containing parental RNA react as sensitive particles (70S couple formation, Fig. 2b, f), whereas those containing mutant RNA behave as resistant particles (no 70S couple formation, Fig. 2d, h). We conclude that the resistance of strain R31 which has an altered 50S subunit is also conferred by the RNA moiety of the large subunit.

Both reconstituted 30S and reconstituted 50S particles containing mutant 16S and (23+5S) RNA, respectively, could form 70S couples with their complementary native subunits only in the presence of higher Mg^{2+} concentration (above 12–15 mM; data not shown). In spite of their normal sedimentation and association behaviour the reconstituted particles were only poorly active in *in vitro* systems of protein synthesis.

Viomycin is the fourth example where resistance against an antibiotic is located in the RNA. Kasugamycin-resistant ribo-

somes contain an altered 16S RNA; the RNA is under-methylated near the 3' end¹². Erythromycin resistance can be due to an over-methylated 23S RNA¹³. Recently, a thiostrepton producing organism was described which protects itself against this drug by methylating the RNA from the large ribosomal subunit¹⁴. However, viomycin is the only case so far known in which the resistance against the drug can be located in the RNA moiety of either the small or the large subunit. Taken together with our observation that viomycin induces association these findings suggest that the RNA of both subunits is involved in the association reaction.

We thank Drs H. Teraoka and Dr. R. Brimacombe for discussions. This work was supported in part by a grant to T.Y. from the Ministry of Education of Japan.

T. YAMADA
Y. MIZUGICHI

Research Institute for Microbial Diseases
Osaka University
Yamada-Kami, Suita City, Osaka, Japan 565

K. H. NIERHAUS
H. G. WITTMANN

Max-Planck-Institute für Molekulare Genetik,
Innistrasse 63-73, Berlin-Dahlem, Germany

Received 24 May; accepted 24 July 1978.

1. Davies, J. E., Gorini, L. & Davis, B. D. *Molec. Pharmacol.* **1**, 93–106 (1965).
2. Tanaka, N. & Igusa, S. *J. Antibiot.* **21**, 239–240 (1968).
3. Masuda, K. & Yamada, T. *Biochim. biophys. Acta* **453**, 333–339 (1976).
4. Yamada, T. & Nierhaus, K. H. *Molec. gen. Genet.* **161**, 261–265 (1978).
5. Lion, Y. F. & Tanaka, N. *Biochem. biophys. Res. Commun.* **71**, 477–483 (1976).
6. Modolell, J. & Vázquez, D. *Eur. J. Biochem.* **81**, 491–497 (1977).
7. Yamada, T., Masuda, K., Shoji, K. & Hori, M. *J. Bact.* **112**, 1–6 (1972).
8. Yamada, T., Masuda, K., Shoji, K. & Hori, M. *Antimicrob. Ag. Chemother.* **6**, 46–53 (1974).
9. Yamada, T., Masuda, K., Mizuguchi, Y. & Suga, K. *Antimicrob. Ag. Chemother.* **9**, 817–823 (1976).
10. Dohme, F. & Nierhaus, K. H. *J. molec. Biol.* **107**, 585–599 (1976).
11. Traub, P., Mizushima, S., Lowry, G. V. & Nomura, M. *Meth. Enzym.* **20**, 391–407 (1971).
12. Helsner, T. L., Davies, J. E. & Dahlberg, J. E. *Nature* **233**, 12–14 (1971); *Nature new Biol.* **235**, 6–9 (1972).
13. Lai, C. J. & Weisblum, B. *Proc. natn. Acad. Sci. U.S.A.* **68**, 856–860 (1971).
14. Cundliff, E. *Nature* **272**, 792–795 (1978).

Corrigendum

In the letter 'Visual response in barnacle photoreceptors is not initiated by transitions to and from metarhodopsin' by Z. Atzmon, P. Hillman and S. Hochstein, *Nature* **274**, 76 (1978), the words 'blue' and 'red' heading the third and fourth columns of Table 1 should be interchanged.

Nature Index and Binders

The complete **Index** for 1977 is available, price £2.50 (U.K.), US\$5.00 (Rest of World). Copies of the 1976 index are still on sale, prices as above.

Binders for the journal are also available at £8.00 (UK), US\$16.00 (Rest of World) for three: a year of *Nature* fits into three binders.

All prices include postage. Payment may be made in any currency at the prevailing exchange rate. Orders should be sent, accompanied by remittance, to Macmillan Journals Ltd, Brunel Road, Basingstoke, Hampshire, RG21 2XS, England.

matters arising

Relative merits of 'extreme-pressure' additives

THE relative merits of 'extreme-pressure' additives are best examined under comparable conditions of 'mild wear' when only a fine dust is abraded from the rubbing surfaces and no macroscopic damage has been done. Macroscopic damage always leads to 'catastrophic' conditions and must therefore be avoided to ensure safe operation for a practically indefinite period. The criterion for safe operation is the validity of Amontons' law. The principal function of any extreme-pressure additive is to extend substantially the range of normal loads in which one can be certain that only mild wear will take place, so that bearing pressures of at least $15,000 \text{ kg wt cm}^{-2}$ are maintained. An optimum concentration¹ exists for each extreme-pressure additive which can react with the lubricated surfaces. Any comparison of extreme-pressure additives must therefore be made at their optimum concentrations and at normal loads for which Amontons' law is obeyed. With 0.9525-cm diameter En-31 steel balls in a four-ball apparatus running at 1,500 r.p.m. the optimum concentrations of two different chlorinated paraffin waxes and of benzyl chloride in a light mineral oil were respectively 5, 2 and 1% by weight, corresponding to 59, 29 and 8 mg-atoms of chlorine per 100 g of oil².

The existence of an optimum concentration suggests that two counteracting processes are responsible for the action of extreme-pressure additives. It is difficult to see how the formation of a protective coating on its own could account for an optimum concentration, but the view expressed recently by McCarroll *et al.*³ that extreme-pressure lubricants containing either sulphur or chlorine form with steel "surface layers of iron sulphide or chloride which prevent metal wear and seizure" is quite unacceptable, because mild wear always takes place, even if the wear prints have a mirror finish. What the extreme-pressure additive does to extend the range of validity of Amontons' law of its base oil, is to lower both the force of friction and the rate of wear which would be encountered with the base oil. At their optimum concentrations each of the chlorine-containing additives mentioned above reduced the coefficient of friction of the base oil by 15%, from 0.100 to 0.085, and extended the range of validity of Amontons' law from 16.5 kg wt to respectively 34, 32 and 26 kg wt (Fig. 1).

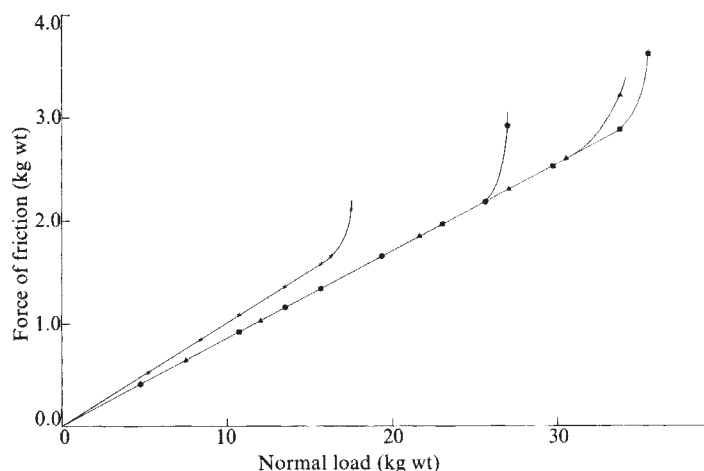


Fig. 1 Force of friction-normal load diagrams (0.9525 cm diameter En-31 steel balls; 1,500 r.p.m.; 30 s runs; additive solutions of optimum concentrations). $\times$, Light mineral oil; $\square$, 5% by wt chlorinated paraffin wax A; $\triangle$, 2% by wt chlorinated paraffin wax B; $\circ$, 1% by wt benzyl chloride (% solutions in light mineral oil).

On this evidence benzyl chloride was less effective than either of the two chlorinated paraffin waxes. The same order of merit was just discernible from the wear scar diameter curves. At 21.6 kg wt for instance, the respective diameters for 30 s runs were 0.330, 0.325, and 0.323 mm. [The bearing pressure of $25,251 \text{ kg wt cm}^{-2}$ at the end of each of these runs was still well within the extreme-pressure range whereas McCarroll and his colleagues had worked so far outside the range of validity of Amontons' law that macroscopic damage had been done and their pressures had fallen to $4,296 \text{ kg wt cm}^{-2}$ and less. Their rate of wear had been nearly 400 times larger than ours.] The corresponding volumes of dust abraded were 1.21×10^{-7} , 1.14×10^{-7} , and $1.11 \times 10^{-7} \text{ cm}^3$. By electron beam induced X-ray analysis considerably

less chlorine was found on wear prints obtained with 1% of benzyl chloride than with the optimum concentration of one of the chlorinated paraffin waxes (Table 1). Wear prints obtained with the base oil alone in the transition region from 'mild' to 'catastrophic wear' (Fig. 2a) resembled those found earlier by Czichos and Kirschke⁴. They had the appearance of a paste having been smeared over the wear print. By contrast, scanning electron micrographs from runs with chlorinated paraffin wax (Fig. 2b) and benzyl chloride (Fig. 2c) in conditions of mild wear showed clearly defined details of the wear scar areas. A striking difference between them was some wear debris sticking to the wear prints obtained with chlorinated paraffin wax (Fig. 2b). These blobs of debris contained more chlorine than the areas around them (Table 1).

Fig. 2 a, SEM photomicrograph ($\times 28,000$) of a wear print on an En-31 steel ball after a 30 s run with a light mineral oil at 15.7 kg wt and 1,500 r.p.m. b, A wear print on an En-31 steel ball after a 30 s run with the optimum concentration (2% by wt) of chlorinated paraffin wax at 21.6 kg wt and 1,500 r.p.m. Blobs of wear debris can be seen on the wear scar. ($\times 2,500$). c, A wear print on an En-31 steel ball after a 30 s run with the optimum concentration (1% by wt) of benzyl chloride at 21.6 kg wt and 1,500 r.p.m. ($\times 2,700$).

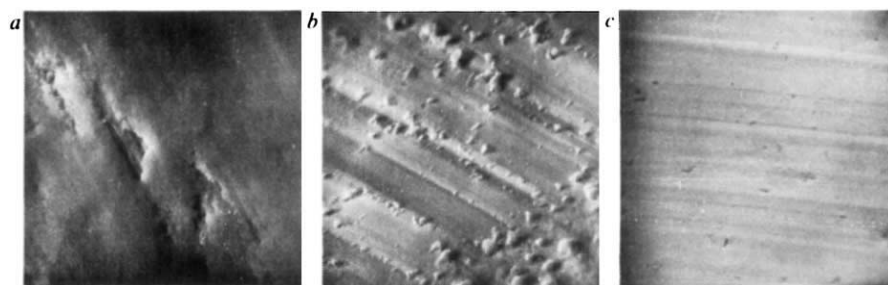


Table 1 Chlorine content found by electron beam induced X-ray analysis of the basis that the contents of Mn, Cr, Fe, and Cl add up to 99.99%

	Percentage of Cl		
	On wear print area	On blobs of debris	On arbitrarily selected spots
5% by wt chlorinated paraffin wax B	0.38	2.48	—
2% by wt chlorinated paraffin wax B (optimum concentration)	0.20	1.34	—
1% by wt benzyl chloride (optimum concentration)	Trace	—	0.10

Wear prints were obtained from 30 s runs with 0.9525-cm diameter En-31 steel balls at 1,500 r.p.m. and 21.6 kg wt.

We conclude from these observations that chlorinated paraffin wax is superior to benzyl chloride as an extreme-pressure additive. While the chemical reactivity of the additive plays some part in the mechanism of extreme-pressure lubrication, it does not lead to the formation of a wear-protective layer.

ROBERT SCHNURMANN
TALAL M. C. PORTER

Department of Chemical Engineering,
University of Birmingham,
Birmingham, UK

1. Schnurmann, R. *J. appl. Phys.* **20**, 376 (1949).
2. Porter, T. M. C. thesis, Birmingham Univ. (1978).
3. McCarroll, J. J., Mould, R. W., Silver, H. B. & Sims, M. L. *Nature* **266**, 518 (1977).
4. Czichos, H. & Kirschke, K. *Wear* **22**, 321 (1972).

SILVER AND MOULD REPLY—We were rather surprised that Schnurmann and Porter found “quite unacceptable” our view¹ that extreme pressure lubricants containing either sulphur or chlorine form with steel surface layers of iron sulphide or chloride which reduce metal wear and prevent seizure. This is not only our view but also that given by Rowe², Prutton³, Davey⁴ and many other workers in the field of extreme pressure lubrication. We were similarly surprised to read the statement that the chemical reactivity of the additive, although playing some part in the mechanism of extreme pressure lubrication, does not lead to the formation of a wear-protective layer. This too is in contrast with our findings⁵, and the findings of others^{6,7}, of the relationship between chemical activity and extreme pressure performance.

However, it is possible that a misunderstanding has arisen over the definition of the lubrication condition known as ‘extreme pressure’. By this we, and others, mean the conditions whereby lubrication is maintained by means of a surface layer of principally inorganic material. These

conditions are operable under the severe loadings found in hypoid gears or in certain metal cutting operations, and are in contrast with those used by Schnurmann and Porter in their experiments. In lower loaded anti-wear tests⁸, we found that chlorinated hydrocarbons gave a slightly better performance than benzyl chloride, although at the same chlorine concentration in each oil; however, we must question the validity of results obtained from blends of chlorine compounds as described by Schnurmann and Porter, each with such different chlorine contents.

H. B. SILVER
R. W. MOULD

BP Research Centre,
Chertsey Road,
Sunbury-on-Thames, UK

1. McCarroll, J. J., Mould, R. W., Silver, H. B. & Sims, M. L. *Nature* **266**, 518 (1977).
2. Rowe, G. W. *RIC Reviews* **135**, 196 (1967).
3. Prutton, C. F., Turnbull, D. & Dlouhy, G. *J. Inst. Petrol.* **32**, 90 (1947).
4. Davy, W. *J. Inst. Petrol.* **31**, 73, 1945.
5. McCarroll, J. J., Mould, R. W., Silver, H. B., Sims, M. L. & Syrett, R. J., *Tribology 1978—Materials, Performance and Conservation* (University of Swansea, 1978).
6. Forbes, E. S., *Wear* **15**, 87 (1970).
7. Rossi, E. & Imparato, L. *Lubrificazione* **4**, 61 (1969).
8. Mould, R. W., Silver, H. B. & Syrett, R. J. *Wear* **22**, 269 (1972).

Interspecific competition in tits

DHONDT has recently made a case for the existence of interspecific competition between blue and great tits¹. My first criticism is mainly statistical and concerns Dhondt's Fig. 3 in which he presents a Lotka–Volterra graph with the equilibrium lines for the two species, estimated subjectively by eight colleagues. I tried a discriminate function analysis as a more rigorous approach and found that the blue tit line had a slight, though insignificant, positive slope. Moreover, in a linear multiple regression of population change of one species on the density of both species, neither species had a significant effect on the other. Consequently, neither the competition coefficients nor the conclusions about niche breadth drawn from the coefficients can be upheld.

My second criticism concerns the interpretation of the significant regressions in Dhondt's Figs 1 and 2. Both figures show that relative to great tits, the reproductive success of blue tits increases as the combined density of the two species increases. Because of the greater variance of the great tit density the result, especially in Fig. 2, could be due to intraspecific, density-dependent effects. In a linear multiple regression of relative competitive ability (great tit fledged per pair divided by blue tit fledged per pair) on year, blue tit density, great tit density, and the ratio of

great tit density to blue tit density, no independent variable is significant although the whole regression is significant. Thus, the dependence of relative competitive ability on combined density seems uncertain and the possible contribution of the blue tit–great tit ratio should not be rejected. Finally, as there will be interspecific differential mortality, emigration, and so on, there is no reason to think that the equilibrium combined density should be where the reproductive abilities of the two species are equal.

British data support the trends reported by Dhondt except that in contrast to Dhondt's study, there is a strong positive correlation between changes in blue and great tit densities².

I thank M. G. Bulmer, J. R. Krebs, and C. M. Perrins for useful discussions.

EDWARD O. MINOT

Edward Grey Institute,
Department of Zoology,
Oxford, UK

1. Dhondt, A. *Nature* **268**, 521–523 (1977).
2. Minot, E. O. (in preparation).

DHONDT REPLIES—In my letter¹ I did two things (1) demonstrate that competition for food in the breeding season between great and blue tit results in a reduced fitness (fewer young produced) and (2) speculate how these titmice nevertheless coexist. Minot confirms point (1) but here criticises point (2).

The first part of Minot's criticism concerns the isoclines I tried to estimate. The method I used was unorthodox. I thought that fitting the lines by eye made it possible to correct for aberrant points caused by varying environmental factors. The more orthodox approach is to calculate a multiple regression equation in which the relative change in numbers (not the absolute value as Minot has done) is expressed as a function of each species' density. Doing this, leaving the 1962 point out (see ref. 1), results in isoclines very similar to the ones drawn in my Fig. 3. These are, for the great tit (GT) $1.817 - 0.075 \text{ GT} - 0.045 \text{ BT} = 0$ and for the blue tit (BT) $1.186 - 0.098 \text{ BT} - 0.0087 \text{ GT} = 0$. It is correct that change in numbers of either species is not significantly correlated with the density of the other species. However both partial correlation coefficients are negative and the one expressing the effect of the BT on GT change in numbers is almost significant ($r = -0.42$, $t_{14} = 1.736$, $P \approx 0.05$, one tailed), although the other one is much lower ($r = -0.23$). We know however, that the key factor, contributing most of the variance to annual mortality is ‘overwinter mortality’^{2,3}. Therefore it would be surprising if change in numbers could be related directly to breeding density in the previous year, without correcting for the factors that cause this variance. I conclude

therefore that the competition coefficients should be treated cautiously but need not be rejected. The conclusions about niche breadth were not, as Minot states, derived from the competition coefficients, but inferred from the results in my Table 1, and only refer to the breeding season. I did not imply that, at other times of the year, this situation holds.

In the second part of his criticism Minot concludes that, because BT breeding numbers vary less than GT numbers, there could be a contribution of the relative proportion of each species to relative competitive ability. This may be so at high combined densities, but using only the data where the variance of the two species is not different (my low combined density years, upper line in Fig. 1) my conclusion is not affected.

I am grateful to Minot for pointing out the difficulty of this type of a posteriori data analysis. I think that critical field experiments designed to test hypotheses will contribute more to our understanding of this problem than further manipulation of the data. The first results from such experiments^{4,5} show that outside the breeding season great and blue tit also compete strongly, but that the situation is reversed and that the blue tit is more affected by it than the great tit.

ANDRÉ A. DHONDT

Departement Biologie,
Universitaire Instelling Antwerpen,
B-2610 Wilrijk, Belgium

1. Dhondt, A. *Nature* **268**, 521–523 (1977).
2. Krebs, J. R. *Ecology* **52**, 2–22 (1971).
3. Dhondt, A. (in preparation).
4. Dhondt, A. A. & Eyckerman, R. (in preparation).
5. Löhrl, H. *Die Vogelwarte Sonderheft*, 92–101 (1977).

Cordilleran Benioff Zones

CONEY AND REYNOLDS¹ have interpreted the radiometrically dated igneous rock pattern in the southwestern boundary of North America in terms of the evolution of a convergent plate margin. They conclude that the pattern reflects a progressive decreasing of the Benioff Zone dip angle from about 130–110 Myr to about 55–40 Myr ago and a rapid increasing of dip from about 40 to 15 Myr. Because of these changes in dip, the magmatic arc zone was displaced about 1,000 km inboard from the convergent margin until 55–40 Myr, then swept it back towards the trench by about 20 Myr. The width of this magmatic zone also changes as a function of the dip angle, from narrow at steep angles to very wide at low angles. Although the data seem to be consistent in supporting their interpretation, another interpretation based on the application of the same plate tectonic scheme^{2–5} used by Coney and Reynolds can be proposed.

According to the model of Dickinson and Hatherton^{2–5}, the surface arc mag-

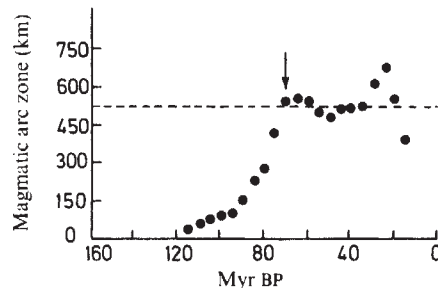


Fig. 1 Growth of the magmatic arc zone. The time when the descending plate approximately reached the 300 km limit of the low-velocity layer is indicated by the arrow.

matism occurs above the zone in which the descending plate intersects the low-velocity layer of the upper mantle. This zone extends from a depth of about 100–300 km. By extrapolating this concept to the earlier stages of development of a magmatic arc, it can be expected that the width of the surface manifestation zone increases as the descending plate reaches greater depths within the low-velocity layer. When the plate reaches the 300 km-limit, if no changes in dip occur, the width of the magmatic arc zone remains nearly constant, until the subduction process ends and the geometry changes. In respect of the zone between 0 and 100 km depth, called the zone of shallow earthquakes, the dip can be different compared to the dip in the low-velocity zone as a consequence of the differences in properties and factors which control the plate interactions in the margin. Thus we can expect two dip angle zones⁶, which have been recognised in several parts^{2–6} and different dip changes in time. At the earlier stages the dip is steep and decreases as the margin evolves until the subduction process finishes or the geometry changes.

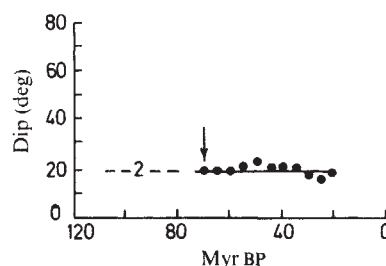


Fig. 2 Magmatic production zone dip angle as a function of time beneath southwestern North America.

By examining the data of Coney and Reynolds, we find that the width of the magmatic zone increases from about 130 to about 70 Myr and remains nearly

constant between 70 and 20 Myr with certain changes from 30 to 20 Myr (Fig. 1). By using a limits of 100 to 300 km for the low-velocity layer and the width of the magmatic arc zone, we can calculate the magmatic production zone dip angle, which is nearly constant from 70 to 20 Myr (Fig. 2). In respect of the zone of shallow earthquakes, the dip angle in this zone experiences a progressive flattening from 130 to about 70–60 Myr, remains constant between 70–60 and 30 Myr, followed by a rapid steepening from 30 to 15 Myr (Fig. 3).

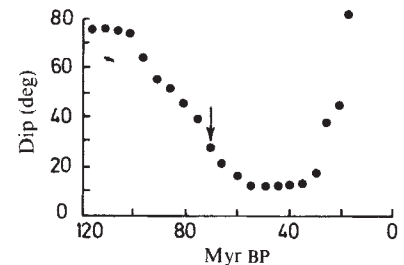


Fig. 3 Shallow earthquakes zone dip angle as a function of time beneath southwestern North America.

The subduction rate during the earlier stages of development can be easily estimated from Fig. 1. The rate increases from 130 to 70 Myr. The changes in the interval between 30 and 20 Myr could represent the evolution of the Kula-Farallon triple junction as suggested in the reconstruction of Atwater⁷. Thus, these changes can be explained as the result of cessation of subduction caused by the change in plates geometry.

It is expected that the model sketched here represents a more reliable approach to understanding the evolution of the southwestern margin of North America. The qualitative aspects of this model were already included in a work under preparation which includes studies of the metallogeny zoning in northwestern Mexico and the evolution of Baja California peninsula. Further details of the present discussion will also be reported elsewhere.

J. URRUTIA FUCUGAUCHI

Department of Geophysics and
Planetary Physics,
School of Physics,
The University,
Newcastle-upon-Tyne, UK

1. Coney, P. J. & Reynolds, S. J. *Nature* **270**, 403–406 (1977).
2. Dickinson, W. R. & Hatherton, T. *Science* **157**, 801–803 (1967).
3. Hatherton, T. & Dickinson, W. R. *J. geophys. Res.* **73**, 4615–4619 (1968).
4. Hatherton, T. & Dickinson, W. R. *J. geophys. Res.* **74**, 5301–5310 (1969).
5. Dickinson, W. R. *Rev. Geophys.* **8**, 813–860 (1970).
6. Urrutia Fucugauchi, J., thesis, Univ. Mexico (1975).
7. Atwater, T. *Geol. Soc. Am. Bull.* **81**, 3513–3536 (1970).

reviews

Our energy alternatives

Walt Patterson

Earth, Water, Wind and Sun: Our Energy Alternatives. By D. S. Halacy Jr. Pp. 186. (Harper and Row: New York and London, 1978.) £4.95.

Too many commentaries on 'alternative energy' seem to assume that we have only just discovered the sun. What is needed is a sense of historical perspective. It is otherwise too easy to overlook the fact that many ingenious technologies to exploit ambient energy flows like wind and sunlight were developed decades, indeed centuries ago. The principles and even the engineering of such technologies are often better understood than those of so-called 'conventional' energy technologies like offshore oil extraction or nuclear power. This historical perspective is one of several distinctive virtues displayed in D. S. Halacy's book, an outstanding example of an under-rated discipline and popular science writing at its best.

Even for a battle-weary reader case-hardened by continual exposure to outpourings about 'alternative energy', Halacy's book has an exhilarating freshness, and is quite unexpectedly absorbing, partly because of the frequently unfamiliar and fascinating detail included and partly because of the writing itself. The book sustains the reader's interest without gimmicks or gee-whizzery, succumbing neither to the breathless flashiness prevalent in the 1950s nor to the more recently fashionable monitory breast-beating and finger-pointing. Furthermore, Halacy reveals himself to be a virtuoso of the short sentence. His prose races along, delivering an astonishing quantity of information accurately and effortlessly, without a hint of condescension.

He sets the stage with an introductory chapter entitled "The Energy Crisis Is Real", or, as he puts it, "But first, the bad news." It would doubtless be possible to quarrel with the picture he paints, at least in some particulars. He is preoccupied almost entirely with energy supply, and discounts far too completely the opportunities for improved efficiency—for optimising the balance between ambient energy, fuel energy, energy conversion systems, and desired end-use objectives. However, he then moves on to devote

chapters to geothermal energy, water-power, tidal power, sea thermal energy, wind, biofuels, and solar energy; and in each case his survey is both crisply concise and impressively comprehensive. He discusses basic principles, pioneering projects, current plans, proposals and prospects for the future, salting the discussion with striking examples and vivid turns of phrase. You probably know about Palmer Putnam's huge windmill on Grandpa's Knob; but did you know about Georges Claude's sea thermal plant with the intake pipe one and one quarter miles long, off the Cuban coast? In 1927? Or consider this thought-provoking aside, on the theoretical limit to recovery of energy from the wind: "(Extracting all the energy would mean bringing all the wind to a dead stop, an interesting possibility.)".

Halacy also avoids a pitfall into which many of his compatriots topple.

Unlike them he does not seem to believe that energy issues stop at the three-mile limit around the US. On the contrary he draws his examples and his ideas from every corner of the globe, although activities in the US do tend to predominate.

The book also offers a wealth of pictorial accompaniment to the text, including some photos of considerable rarity value.

All told, in 186 pages Halacy has put together a book which can be recommended with enthusiasm as one of the best available introductions to "our energy alternatives". Even its upbeat final sentence is welcome: "It's our move—and the prospects are encouraging". The next time some official spokesman threatens you with freezing in the dark, introduce him to Halacy.

Walt Patterson is energy consultant to Friends of the Earth (London) Ltd.

Ecology of freshwater macrophytes

River Plants: The Macrophytic Vegetation of Watercourses. By S. M. Haslam. Pp. 396. (Cambridge University Press: Cambridge and London, 1978.) £27.50.

AGNES ARBER published her outstanding book on water plants nearly 60 years ago. Apart from Fritz Gessner's *Hydrobotanik*, Dr Haslam's is the first dealing with these plants, and based on original research, to have appeared since then. A causal analysis of the distribution of river plants throughout Britain is attempted in terms of flow, substrate and nutrients. These findings are applied to downstream changes in habitat and species composition in whole rivers. This leads to a classification of the vegetation of streams on soft rocks, streams on hard rocks and, finally, of channels with little flow. A preliminary account follows of the vegetation of watercourses in eastern North America and the book concludes with five chapters on applied aspects—uses and hazards created by river plants, their management and their relationship to pollution.

The author clearly intends to interest and indeed educate non-botanists and particularly those engaged in water

planning and management; there is an initial set of drawings of 'field' plants, and symbols are used in text, tables and river maps for the commonest species. Perhaps this explains the way degrees of correlation between nutrient and species distribution are presented (Tables 8.2, 8.6) and the grouping of all moss data as 'mosses'. The professional biologist or teacher may find it difficult to extract succinct information because of inevitable compression of data and, sometimes, insufficient distinction between generalisations based on these data and on subjective assessment. The book is very well produced, with good illustrations, pleasant lay-out and few errors—the most serious being the omission of all data from Fig. 6.12. These minor faults only slightly mar an outstanding, single-handed contribution to river biology—the first classification in terms of their macrophytes of British (and some North American) watercourses. The book can confidently be recommended to all those interested in the ecology of freshwaters.

D. H. N. Spence

D. H. N. Spence is Professor and Head of the Department of Botany at the University of St Andrews, UK.

Origins of senescence

The Genetics of Aging. Edited by E. L. Schneider. Pp.424. (Plenum: New York and London, 1978.) \$39.

DESPITE a considerable amount of research in recent years, as one of the contributors to this volume states, "... a universally acceptable definition of normal biological aging is still not available. The origins of senescence remain enigmatic, and little agreement exists regarding its true nature". Having read this well referenced up-to-date review of the subject one cannot help but agree with this sentiment, but as the nature and cause of aging raise basic fundamental questions in biology perhaps this is not too surprising.

The book is divided into four parts. The first is concerned with the possible molecular basis of aging which resolves essentially into error and programmed theories. The former assumes that aging results from the accumulation of errors in DNA replication or deficiencies in the repair of such errors. The latter assumes that aging results from genetically programmed senescence. Evidence is presented both for and against these two possibilities. The observation in man of increasing hypodiploidy in cultured lymphocytes with age, most frequently affecting the X chromosome in females and the Y chromosome in males, is intriguing, but whether this is a cause or effect of aging remains unresolved.

The second part of the book is concerned with aging at the level of the organism itself from protozoa and fungi to vertebrates. Whether certain findings in lower organisms are of general relevance is a mute point. For example, in *Podospora anserina* (a filamentous fungus) there is clear evidence that longevity is at least partly controlled by cytoplasmic factors. In higher forms life span and survivorship can be expressed in terms of allometric and Gompertz equations for which Sacher presents convincing explanations. It would seem reasonable, however, that aging in man is at least partly determined by genetic factors and the third and largest part of the book is devoted to this subject. In 1961 Hayflick first demonstrated that human diploid somatic cells have a limited replicative life span in tissue culture which is inversely related to the age of the donor. Since then, this has become an exciting and popular experimental tool for studying the phenomenon of aging at least as this is expressed *in vitro*. Detailed consideration is given to this phenomenon in the case of those genetic syndromes known

to be associated with premature aging such as progeria in which fibroblasts in culture show reduced survival, defective HLA expression, increased heat lability of cytoplasmic enzymes, changes in insulin receptors and tissue factors promoting clotting. These observations are appealing; but are such disorders a legitimate paradigm of normal aging? The possible role of genetic factors in aging is more directly approached by family and twin studies. The findings in the now famous Pearl (Baltimore) study of longevity among relatives of nonagenarians are discussed in detail and indicate that there is a definite familial component in aging. However, this is small and Murphy is careful to avoid being dogmatic as to its nature. It could well be largely cultural or environmental. On the other hand, the discussion of

twin studies gives more emphasis to a genetic component in aging.

The fourth and final part of the book is concerned with certain experimental approaches to aging, including somatic cell genetics, immunogenetics and behavioural genetics.

Inevitably in a work of this nature there is some repetition but this is less than in most large multi-authored books. In any event it is often interesting to compare the views of different contributors on the same subject; though the index itself is often unhelpful in this regard. This is a highly readable and well referenced review of a fascinating subject and I would recommend it without reservation.

Alan E. H. Emery

Alan E. H. Emery is Professor of Human Genetics at the University of Edinburgh, UK.

Nuclear transplantation experiments in Amphibia

Cloning: Nuclear Transplantation in Amphibia. By R. G. McKinnell. (University of Minnesota Press: Minneapolis, 1978.) \$22.50.

IN spite of its principal title, this book is not concerned with most aspects of cloning. Thus, it does not deal with cloning in plants, with the cloning of DNA or genes in bacteria by plasmids and phages, nor with the cloning of animal cells in culture. In fact the book consists entirely of an account of nuclear transplantation experiments in Amphibia, primarily those with *Rana pipiens*. This has its use, and a chapter on the transplantation of nuclei from tumour cells is clearly presented and well balanced.

Unfortunately the same cannot be said for much of the rest of the book. In general, work with species other than *R. pipiens* is treated as trivial or at best confirmatory. Often it is hard to discern the direction of an argument. Thus, the book does not make it clear that many amphibian nuclear transfer experiments have been addressed to the specific question of whether the nucleus of one kind of specialised cell contains the genetic information necessary to promote the formation of other quite unrelated cell types. Though generally considered important, this question is totally obscured in this book by the author's decision to emphasise the age rather than differentiated state of donor cells.

From the student's point of view, a

serious deficiency is that alternative but not exclusive interpretations are often omitted. For example, most of the first half of the book is concerned with abnormal embryos resulting from nuclear transfers in *R. pipiens*. It is therefore disappointing to find that one of the principal ideas regarding the significance of these abnormal embryos is entirely overlooked. The author presses the view that the abnormalities of nuclear transplant embryos may give specific information about the specialised state of gene control in donor cells. An alternative interpretation (favoured by several in this field) is that abnormal embryos result from random chromosomal losses due to the low replication rate of specialised cells and the very high rate at which eggs always divide. My criticism is not that the author subscribes to the former point of view, but to disguise or ignore the second interpretation seems to me to significantly misrepresent the literature.

The last half of the book contains useful technical information in the form of six appendices. These deal with the biology and maintenance of *R. pipiens*, the instruments, solutions and techniques used to transplant single nuclei to eggs (primarily *R. pipiens*), as well as with sources of various anuran species.

The unbalanced presentation and uncritical interpretation of nuclear transplantation experiments seriously limit the general usefulness of the first part of this book. However, the technical information in the appendices should prove valuable for those wishing to transplant nuclei in *R. pipiens*.

J. B. Gurdon

J. B. Gurdon is at the MRC Laboratory of Molecular Biology in Cambridge, UK.